

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 13:50:30 ; Search time 8904.87 Seconds
(without alignments)
2958.109 Million cell updates/sec

Title: US-08-894-356C-1
Perfect score: 1703
Sequence: 1 TCATTATGGAGCAATCCAA.....AAAAAAAAAAAAAAAAAAAA 1703

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*

1:	gb_bal:*
2:	gb_ba2:*
3:	gb_ba3:*
4:	gb_in1:*
5:	gb_in2:*
6:	gb_in3:*
7:	gb_om:*
8:	gb_ov:*
9:	gb_pat1:*
10:	gb_pat2:*
11:	gb_ph:*
12:	gb_pl1:*
13:	gb_pl2:*
14:	gb_pl3:*
15:	gb_pl4:*
16:	em_bal:*
17:	em_ba2:*
18:	em_fun:*
19:	em_htgo_hum:*
20:	em_htgo_inv:*
21:	em_htgo_rod:*
22:	em_htg_hum1:*
23:	em_htg_hum2:*
24:	em_htg_hum3:*
25:	em_htg_hum4:*
26:	em_htg_hum5:*
27:	em_htg_hum6:*
28:	em_htg_hum7:*
29:	em_htg_hum8:*
30:	em_htg_inv1:*
31:	em_htg_inv2:*
32:	em_htg_other:*
33:	em_htg_rod:*
34:	em_hum1:*
35:	em_hum2:*
36:	em_hum3:*
37:	em_hum4:*
38:	em_hum5:*
39:	em_hum6:*
40:	em_hum7:*
41:	em_in:*
42:	em_om:*
43:	em_or:*

44:	em_ov:*
45:	em_pat:*
46:	em_ph:*
47:	em_pl:*
48:	em_ro:*
49:	em_sts:*
50:	em_sy:*
51:	em_un:*
52:	em_vl:*
53:	gb_sts1:*
54:	gb_sts2:*
55:	gb_sts3:*
56:	gb_sy:*
57:	gb_un:*
58:	gb_vl1:*
59:	gb_vl2:*
60:	gb_htg1:*
61:	gb_htg2:*
62:	gb_htg3:*
63:	gb_htg4:*
64:	gb_htg5:*
65:	gb_htg6:*
66:	gb_htg7:*
67:	gb_htg8:*
68:	gb_htg9:*
69:	gb_htg10:*
70:	gb_htg11:*
71:	gb_htg12:*
72:	gb_htg13:*
73:	gb_htg14:*
74:	gb_htg15:*
75:	gb_htg16:*
76:	gb_htg17:*
77:	gb_htg18:*
78:	gb_htg19:*
79:	gb_htg20:*
80:	gb_htg21:*
81:	gb_htg22:*
82:	gb_htg23:*
83:	gb_htg24:*
84:	gb_htg25:*
85:	gb_pr1:*
86:	gb_pr2:*
87:	gb_pr3:*
88:	gb_pr4:*
89:	gb_pr5:*
90:	gb_pr6:*
91:	gb_pr7:*
92:	gb_pr8:*
93:	gb_pr9:*
94:	gb_r01:*
95:	gb_r02:*
96:	gb_in4:*
97:	gb_pr10:*
98:	em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1703	100.0	1703	10	E12753
2	1679	98.6	1679	12	AB010708 Gentiana
3	204	12.0	1508	10	E12757 Senecio cru
4	189.2	11.1	1622	10	E12754 Gentianatri
5	189.2	11.1	1622	12	AB026494 Gentiana
6	172.2	10.1	1479	10	E12756 Perilla oci
7	171.4	10.1	1476	12	AB029340 Perilla f
8	114	6.7	11914	12	AC003027 Arabidops

9	109.2	6.4	101176	12	AC002560	Genomic s
10	81	4.8	1648	9	AX025514	Sequence
c 11	71.8	4.2	82360	13	AP000606	Arabidops
c 12	71.8	4.2	341064	75	AC074226	Arabidops
c 13	64.2	3.8	66237	12	AB016892	Arabidops
c 14	63	3.7	136047	74	AC069470	Arabidops
c 15	61.8	3.6	85690	12	AB028618	Arabidops
c 16	61.8	3.6	136047	74	AC069470	Arabidops
c 17	51.2	3.0	71623	77	AC087724	Arabidops
c 18	51.2	3.0	85690	12	AB028618	Arabidops
c 19	49	2.9	1141	10	AX083744	Sequence
c 20	45.6	2.7	321003	84	PF0414P3	Sequence
c 21	44.8	2.6	192929	60	AC005505	Plasmodiu
c 22	44.4	2.6	269	6	ECA270224	Entodiniu
c 23	44	2.6	46739	13	AP002052	Arabidops
c 24	44	2.6	78379	12	AB006696	Arabidops
c 25	43.8	2.6	191166	80	AL356295	Homo sapi
c 26	43.2	2.5	1141	10	AX083744	Sequence
c 27	42.8	2.5	3314	5	AF163834	Dictyoste
c 28	42.8	2.5	188925	67	AC022050	Homo sapi
c 29	42.4	2.5	129838	75	AC079038	Oryza sat
c 30	42.4	2.5	154425	60	AC008911	Homo sapi
c 31	42.4	2.5	156060	60	AC004153	Plasmodiu
c 32	42.2	2.5	186974	83	AP003071	Homo sapi
c 33	42	2.5	156060	60	AC004153	Plasmodiu
c 34	42	2.5	170921	83	CNS01DXJ	Homo sapi
c 35	42	2.5	202487	83	CNS01DW6	Homo sapi
c 36	42	2.5	202496	91	CNS01DX6	Human chr
c 37	42	2.5	219936	64	AC016299	Homo sapi
c 38	41.8	2.5	2222	9	AR105363	Sequence
c 39	41.8	2.5	2222	97	HUMKEINF	Homo sapiens
c 40	41.8	2.5	197120	79	AL354820	Homo sapi
c 41	41.8	2.5	200543	64	AC016726	Homo sapi
c 42	41.6	2.4	356	14	PSN06DMR	P. sativum P
c 43	41.2	2.4	14955	5	AF000580	Dictyoste
c 44	41.2	2.4	100000	91	AP000096	Homo sapi
c 45	41.2	2.4	100000	91	AP000200	Homo sapi

ALIGNMENTS

RESULT	1
E12753	
LOCUS	E12753 1703 bp DNA
DEFINITION	Gentianatriflora mRNA acyltransferase, complete cds.
ACCESSION	E12753
VERSION	E12753.1 GI:3251585
KEYWORDS	JP 1997070290-A/1.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 1703)
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL	Patent: JP 1997070290-A 1 18-MAR-1997; SUNTORY LTD
COMMENT	OS Gentianatriflora PN JP 1997070290-A/1 PD 18-MAR-1997 PR 30-JAN-1996 JP 1996046534 PF 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI PI ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO. PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC C12N15/09, A01H1/00, C07H21/04, C07K14/42, C12N9/10, (C12N9/10, PC C12R1:865), PC (C12N9/10, C12R1:19); CC strandedness: Double; CC topology: Linear; FH key Location/Qualifiers


```
QY 781 AAATCAGAGAGCTAAAGACAAAGTACTGAATCTCAGAGATCCGAAACCGACAATACGTG 840
Db 781 AAATCAGAGAGCTAAAGACAAAGTACTGAATCTCAGAGATCCGAAACCGACAATACGTG 840
QY 841 TAACGACGTTCAATGACGTGGATACGTATGACATGCGATGTCGAATCAAAAGATG 900
Db 841 TAACGACGTTCAATGACGTGGATACGTATGACATGCGATGTCGAATCAAAAGATG 900
QY 901 AGTCTGATCAGAGGAATCATCAACGACGAAAAATGAGTTCGAGTACTTTCAGTACAG 960
Db 901 AGTCTGATCAGAGGAATCATCAACGACGAAAAATGAGTTCGAGTACTTTCAGTACAG 960
QY 961 CGGATTGCCGAGGACTTCTGACGCCGCCCTGTCGCCCTAACTACTTTGCGCAACTGCTTG 1020
Db 961 CGGATTGCCGAGGACTTCTGACGCCGCCCTGTCGCCCTAACTACTTTGCGCAACTGCTTG 1020
QY 1021 CGTCATGCGTTGCAAAAGCAACACATAAAGAGTTAGTTGCGGATAAAGGCTTCTTGTTG 1080
Db 1021 CGTCATGCGTTGCAAAAGCAACACATAAAGAGTTAGTTGCGGATAAAGGCTTCTTGTTG 1080
QY 1081 CAGTTCCAGCTATTGGAGAGCCATTGAAAGAGGTTGCACACGAAAAAGGCTTCTTG 1140
Db 1081 CAGTTCCAGCTATTGGAGAGCCATTGAAAGAGGTTGCACACGAAAAAGGCTTCTTG 1140
QY 1141 CAGATGCAAAAACTTGGTTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTCTCGGGA 1200
Db 1141 CAGATGCAAAAACTTGGTTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTCTCGGGA 1200
QY 1201 TTACCGGATCGCCTAAGTTGCGATTGCGTATGTTGTTAGATTTTGATGGGGAAGCCCTGCAA 1260
Db 1201 TTACCGGATCGCCTAAGTTGCGATTGCGTATGTTGTTAGATTTTGATGGGGAAGCCCTGCAA 1260
QY 1261 AATTTGACATTTACCTCTGTTGATATGACAGAAATGATTTATGATGATTCAGTCCAGGAT 1320
Db 1261 AATTTGACATTTACCTCTGTTGATATGACAGAAATGATTTATGATGATTCAGTCCAGGAT 1320
QY 1321 TTGAAAAAGGTGGAGATGGAGTATCATGCTAAGATTCATATGATGATGATTTGCAAA 1380
Db 1321 TTGAAAAAGGTGGAGATGGAGTATCATGCTAAGATTCATATGATGATGATTTGCAAA 1380
QY 1381 AAATCTTTGAAGAGGCTTTGCTCTTTGTCATAGTCTCTTTAATAGAACCATATTTGCT 1440
Db 1381 AAATCTTTGAAGAGGCTTTGCTCTTTGTCATAGTCTCTTTAATAGAACCATATTTGCT 1440
QY 1441 GCATTAAGTACCAAGTCTTTAGTAACACTACACCAAAACCCCTACTTTTCGAGCGGGAAC 1500
Db 1441 GCATTAAGTACCAAGTCTTTAGTAACACTACACCAAAACCCCTACTTTTCGAGCGGGAAC 1500
QY 1501 ACCAACGAGGTTCAATCACTAGAGGTTGTACTTTATATTAATTCAGAGGTCGAATATA 1560
Db 1501 ACCAACGAGGTTCAATCACTAGAGGTTGTACTTTATATTAATTCAGAGGTCGAATATA 1560
QY 1561 CACCGTTGCTCTGAAAAGTTGAACCTCACACCTGACATGTTGTTACGATAGGTTGTT 1620
Db 1561 CACCGTTGCTCTGAAAAGTTGAACCTCACACCTGACATGTTGTTACGATAGGTTGTT 1620
QY 1621 ATATGCCATATATATCTTCCATAAAGTATCCTATGCAATAGAGAATGTTATGTTGTTA 1680
Db 1621 ATATGCCATATATATCTTCCATAAAGTATCCTATGCAATAGAGAATGTTATGTTGTTA 1680
QY 1681 AAAAAAAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 1703
Db 1681 AAAAAAAGGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTA 1703

RESULT 2
AB010708 1679 bp mRNA PLN 20-FEB-1999
LOCUS Gentiana triflora mRNA for Anthocyanin 5-aromatic acyltransferase,
DEFINITION complete cds.
ACCESSION AB010708
VERSION AB010708.1 GI:4185598
KEYWORDS Anthocyanin 5-aromatic acyltransferase.
```

```
SOURCE Gentiana triflora petal cDNA to mRNA, clone:pgAT4.
ORGANISM Gentiana triflora
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
REFERENCE 1 (bases 1 to 1679)
AUTHORS Fujiwara,H., Tanaka,Y. and Kusumi,T.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Hiroyuki Fujiwara, Suntory LTD., Inst. of Fundamental Research;
1-1-1, Wakayamadai, Shamamoto-cho, Osaka 618-0024, Japan
(E-mail:Hiroyuki_Fujiwara@suntory.co.jp, Tel:+81-75-962-8807,
Fax:+81-75-962-8262)
REFERENCE 2 (sites)
AUTHORS Fujiwara,H., Tanaka,Y., Yonekura-Sakakibara,K.,
Fukuchi-Mizutani,M., Nakao,M., Fukui,Y., Yamaguchi,M., Ashikari,T.
and Kusumi,T.
TITLE cDNA cloning, gene expression and subcellular localization of
anthocyanin 5-aromatic acyltransferase from Gentiana triflora
JOURNAL Plant J. 16 (4), 421-431 (1998)
MEDLINE 99097837
FEATURES Location/Qualifiers
source 1..1679
/organism="Gentiana triflora"
/db_xref="taxon:55190"
/clone="pgAT4"
/tissue.type="petal"
CDS 6..1415
/codon_start=1
/product="Anthocyanin 5-aromatic acyltransferase"
/protein_id="BAA74428.1"
/db_xref="GI:4185598"
/translation="MEQIQWKVLEKCOVTPSDTDDVELSLPVTFFDIPWLHNKMQ
SLLEFDYPPRTHFLDTVLPNLKASLTLKHVPLSGNLMLPIKSGEMPKQYSRDE
GDSITLIVAE:SDQDEFYLGKQLVDSNDLHGLFYVMPRVIRTMQDYKVIPLVAVQTV
FPNRGIAVALTAHSHDAKSEFMFINAKGKADLLSFLDSFDRSIKDL
YGLETFPWNEMODVLEMFRRFGSKPPFNKVRATYVLSLAIEIQRKLNKVLNURGSEPT
IRVTTFTMTGTYVTKYKSKDDVYSEESNDENELEYFSETADRGILITPPCPNYF
GNCLASCAKATKBLVGDGLLVAAGAIEAKRHNKGVLAADAKTWLSENGIP
SKRFLGITGSPKFDYGVDFGKPKAKFDITSDVYAEIIYVYIQSRDFEKGVEIGVSLP
KIHMDFAPAKIFEFGFCSLS"
polyA_site 1679
/note="24 a nucleotides"
BASE COUNT 488 a 350 c 359 g 482 t
ORIGIN
Query Match 98.6%; Score 1679; DB 12; Length 1679;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1679; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 TCATTATGAGCAAAATCCAAATGGTGAAGTTCCTTCAAAAATGCCAAGTTACACCACCAT 60
Db 1 TCATTATGAGCAAAATCCAAATGGTGAAGTTCCTTCAAAAATGCCAAGTTACACCACCAT 60
QY 61 CTGACACAACAGATGTCGAGTTATCGCTACCGGTAACATTTCTCGATATCCCTGGTTGC 120
Db 61 CTGACACAACAGATGTCGAGTTATCGCTACCGGTAACATTTCTCGATATCCCTGGTTGC 120
QY 121 ACTTGAATAGATGCAAGTCCCTTCTGTTTACGACTTTTCGTAACCAAGAACACATTTCT 180
Db 121 ACTTGAATAGATGCAAGTCCCTTCTGTTTACGACTTTTCGTAACCAAGAACACATTTCT 180
QY 181 TGGACACTGTTATCCCTTAATCTTAAGCCCTCTTGTCTCTCACTCTAAACACTAGCTTC 240
Db 181 TGGACACTGTTATCCCTTAATCTTAAGCCCTCTTGTCTCTCACTCTAAACACTAGCTTC 240
QY 241 CGCTTAGCGAAATTTGTTGATGCCGATCAAAATCGCGGCAAAATGCCGAAGTTTCAGTACT 300
Db 241 CGCTTAGCGAAATTTGTTGATGCCGATCAAAATCGCGGCAAAATGCCGAAGTTTCAGTACT 300
QY 301 CCCGTGATGAGGCGGACTCCATAACTTTGATCGTTCCGGAGTCTGACCAGGATTTTGACT 360
Db 301 CCCGTGATGAGGCGGACTCCATAACTTTGATCGTTCCGGAGTCTGACCAGGATTTTGACT 360
```

Db	301	CCCGTGATGAGGGCGACTCGATAACTTTTGATCGTTTGGCGAGTCTGACCAAGATTTTGACT	360
Qy	361	ACCTTAAAGGTCAATCAACTGGTAGATTCCAATGATTTGCTAGCCCTTTTATGTATGTC	420
Db	361	ACCTTAAAGGTCAATCAACTGGTAGATTCCAATGATTTGATGGCCCTTTTATGTATGTC	420
Qy	421	CACGGGTTATPAGGACCATTGCAAGACTATAAAGTGTATCCCGTCTGAGCCGTGCAAGTAA	480
Db	421	CACGGGTTATPAGGACCATTGCAAGACTATAAAGTGTATCCCGTCTGAGCCGTGCAAGTAA	480
Qy	481	CCGTTTTTCTAACCGTGGCATAGCCGTGGCTCTGACGGCAGATCATTCATTTGCAGATG	540
Db	481	CCGTTTTTCTAACCGTGGCATAGCCGTGGCTCTGACGGCAGATCATTCATTTGCAGATG	540
Qy	541	CTAAAGTTTTGTAAATGTTTCATCAATGCTTGGCCCTATATTAAACAAATTCGGGAAAGACG	600
Db	541	CTAAAGTTTTGTAAATGTTTCATCAATGCTTGGCCCTATATTAAACAAATTCGGGAAAGACG	600
Qy	601	CGGACTTGTGTCGCCGAATCTTCTTCCATCTTTCGATAGATCGATAATCAAAGATCTGT	660
Db	601	CGGACTTGTGTCGCCGAATCTTCTTCCATCTTTCGATAGATCGATAATCAAAGATCTGT	660
Qy	661	ATGGCCTAGAGGAAACATTTTGGACGAATGCGAAGATGTTCTTGAATGTTCTCTAGAT	720
Db	661	ATGGCCTAGAGGAAACATTTTGGACGAATGCGAAGATGTTCTTGAATGTTCTCTAGAT	720
Qy	721	TTGGAAGCAAAACCCCTCGATTCAACAAAGTACGAGCTACATATGCTCTCCCTTGGCTG	780
Db	721	TTGGAAGCAAAACCCCTCGATTCAACAAAGTACGAGCTACATATGCTCTCCCTTGGCTG	780
Qy	781	AAATCCAGAAAGCTTAAAGAAACAAAGTACTGAAATCTCAGAGGATCCGAACGCAATAGCTG	840
Db	781	AAATCCAGAAAGCTTAAAGAAACAAAGTACTGAAATCTCAGAGGATCCGAACGCAATAGCTG	840
Qy	841	TAACGAGGTTACAAATGACGTGTGGATACGTTATGGACATGCATAGTCAAATCAAAGATG	900
Db	841	TAACGAGGTTACAAATGACGTGTGGATACGTTATGGACATGCATAGTCAAATCAAAGATG	900
Qy	901	ACGTCTGTATCAGAGGAATCATCGAAGCAGCAAAATGAGCTCGAGTACTTTCAGTTTACAG	960
Db	901	ACGTCTGTATCAGAGGAATCATCGAAGCAGCAAAATGAGCTCGAGTACTTTCAGTTTACAG	960
Qy	961	CGGATGCCGAGGACTCTTGAGCCGCCCTGTGCCGCTAACTACTTTGGCAACTGTCTTG	1020
Db	961	CGGATGCCGAGGACTCTTGAGCCGCCCTGTGCCGCTAACTACTTTGGCAACTGTCTTG	1020
Qy	1021	CGTCATCGTTGCAAAAGCAACATATAAGATTTAGTTGGGATAAAGGGCTCTCTGTG	1080
Db	1021	CGTCATCGTTGCAAAAGCAACATATAAGATTTAGTTGGGATAAAGGGCTCTCTGTG	1080
Qy	1081	CAGTTGCAGCTATTGGAGAAGCCATTGAAAGAGGTTGCACACGAAAAGGCGTCTCTG	1140
Db	1081	CAGTTGCAGCTATTGGAGAAGCCATTGAAAGAGGTTGCACACGAAAAGGCGTCTCTG	1140
Qy	1141	CAGATGCAAAACTTGGTTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTCTCGGGA	1200
Db	1141	CAGATGCAAAACTTGGTTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTCTCGGGA	1200
Qy	1201	TTACCGGATCGCCTAAGTTTCGATTCGATATGTTGATTTGGATGGGAAAGCCTGCAA	1260
Db	1201	TTACCGGATCGCCTAAGTTTCGATTCGATATGTTGATTTGGATGGGAAAGCCTGCAA	1260
Qy	1261	AATTTGACATTAACCTCTGTGATTATGCAGAAATTCATTTATGTGATTCAGTCCAGGGATT	1320
Db	1261	AATTTGACATTAACCTCTGTGATTATGCAGAAATTCATTTATGTGATTCAGTCCAGGGATT	1320
Qy	1321	TTGAAAAAGGTGTGGAGATTGGAGTATCATTCGCCCTAAGATTCATATGATGCATTTGCCAA	1380
Db	1321	TTGAAAAAGGTGTGGAGATTGGAGTATCATTCGCCCTAAGATTCATATGATGCATTTGCCAA	1380
Qy	1381	AAATCTTTGAAGAGGCTTTTGCTCTTTGCTCATAGTCTCTTTTAAATAGAACCATATTTGCT	1440
Db	1381	AAATCTTTGAAGAGGCTTTTGCTCTTTGCTCATAGTCTCTTTTAAATAGAACCATATTTGCT	1440

Qy	1441	GCATAAAGTACCAAGTCTTTAGTAAACACTACACAAACCCTACTTTTCGAGCGGGAAC	1500
Db	1441	GCATAAAGTACCAAGTCTTTAGTAAACACTACACAAACCCTACTTTTCGAGCGGGAAC	1500
Qy	1501	ACCACAAGGAGTTCAATCACTAGAGGTTGTACTTCTATAAATTCAGAGGTCGAATATA	1560
Db	1501	ACCACAAGGAGTTCAATCACTAGAGGTTGTACTTCTATAAATTCAGAGGTCGAATATA	1560
Qy	1561	CACGGTCTCCTCGAAAAGTTGAACCTTCACACCTGCACATGGTTCAGCATAGGTATTGT	1620
Db	1561	CACGGTCTCCTCGAAAAGTTGAACCTTCACACCTGCACATGGTTCAGCATAGGTATTGT	1620
Qy	1621	ATAATGCCATTATATACATCCCAATGAAGTATCCTATGCAATAGAGAACATGTTATGTGTT	1679
Db	1621	ATAATGCCATTATATACATCCCAATGAAGTATCCTATGCAATAGAGAACATGTTATGTGTT	1679
RESULT	3		
LOCUS	E12757	1508 bp	DNA
DEFINITION	Senecio cruentus mRNA for acyltransferase,,partial cds.	PAT	24-JUN-1998
ACCESSION	E12757		
VERSION	E12757.1	GI:3251589	
KEYWORDS	JP 1997070290-A/5.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1508)		
AUTHORS	Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y., Yonekura,K., Mizutani,M. and Kusumi,T.		
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY		
JOURNAL	Patent: JP 1997070290-A 5 18-MAR-1997;		
COMMENT	SUNTORY LTD OS Senecio cruentus PN JP 1997070290-A/5 PD 18-MAR-1997 PF 30-JAN-1996 JP 1996046534 PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHITO, PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PI C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC C12R1:865), PC (C12N9/10,C12R1:19); CC strandedness: Double; CC topology: Linear; FH Key Location/Qualifiers FH FT source 1..1508 FT /organism='Senecio cruentus' FT /tissue_type='petal', FT /clone='pCAT48', FT CDS 1..1367 FT /product='acyltransferase'. FT Location/Qualifiers 1..1508 /organism='unidentified' /db_xref='taxon:32644' BASE COUNT 442 a 293 c 296 g 477 t ORIGIN		
Query Match	12.0%	Score 204;	DB 10; Length 1508;
Best Local Similarity	52.1%;	Pred. No. 1.6e-41;	
Matches 721;	Conservative 0;	Mismatches 615;	Indels 48; Gaps 10;
Qy	25	TGAAGTCTTCGAAAAATGCCAAGTTACACCACCATCTGCACACACACATGTCGAGTTAT	84
Db	1	TGAACATTTCTCGAACATCCCGGAATATCGGCCCTCGGGCACCACCA---TCGGCCATCGCT	57
Qy	85	CGCTACCCGGTAAACATTCCTTCGATATCCCTCGTTCACCTTGAATAAGATGCAGTCCCTTC	144

Dbb 58 CGTTATCTCTTACTTTCTTCGACATTACTTGGCTACTCTTCCCTCGGGTCCACCATCTTT 117
QY 145 TGTCTTTACGACTTTCGGTACCCCAAGAACACATTTCTTGGACACTGTATCCCTTAATCTTA 204
Dbb 118 TCTTCTATGACTTTCACATCTTAATCCATTTTCATGGACACTATTTGTTCCCGAGGCTAA 177
QY 205 AGCCCTCTTTGCTCTCACTCTAAAAACACTACCTTCCCGTTAGCGGAAATTTTGTGATGC 264
Dbb 178 AACAACTCTTTATCGGTCACCTCTTCAACATTTTTCCTCGTTGCTAGTAATTTGATGTAT 237
QY 265 CGATCAA-----ATCGGGCGAAATGCCGAAGTTTTCAGTACTCCCGTGTAGAGGGG 315
Dbb 238 TTCTTAACACTGATGTTGGGTTTAAATAAAAAACAGAAATFAAACACGTTTGAAGGTG 297
QY 316 ACTCGATAACTTTGATCGTTGCGGAGCTGACCCAGGATTTTGTACTACCTTAAAGGTCATC 375
Dbb 298 ATTCTGTTGGGTTACTTTTGCAGAAATGTTGCTTGCATTTAATAATTTGCACAGGAATC 357
QY 376 AACTGGTAGATTCCTAATGATTTGCATGGCTTTTATGCTTATGCCACGGTTTATAAGGA 435
Dbb 358 ATCCTCGAAATGTGAAACTTTTATCCACTTGTACTTTCATTG---GGAATGCAATCA 414
QY 436 CCATGCAAGACTATAAGTATGCCCGTGTAGCCGTGCAAGTAAACGGTTTTCCTTAACC 495
Dbb 415 AATTATGTGATGCGTCACGGTCCCACTTTTTCACTTCAAGTGACGTTTTCCTCGGGCT 474
QY 496 GTGCCATAGCGGTGCTGACGGCACATCTCAATTCGAGATGCTAAAAGTTTGTAA 555
Dbb 475 CGGATATATCACTAGGAATGACCAATCATCATAGCCCTTGGTGACGTAGCACCGGTTCA 534
QY 556 TGTTCATCAATGCTTGGGCTTATTAACAAATTTGGGAAAGACGGGACTTCTTGTCCG 615
Dbb 535 ACTTTTGAAGGTCGACTTCCGATTATCAATCTGCTGATAGCTCGTCTTTTAAACGA 594
QY 616 CGAATCTTCTTCATCTTTCGATAGATGATCAATCAAGATCTGTATGGCCTAGAGAA 675
Dbb 595 AAGGATCTCCACCGGTTTGTATGATGATGATTAAACA-----TCCACATTTTATGTA 646
QY 676 CATTTTGGAAAGAAATGCAAGATGTTCTTGAATGTTCTCTAGATTTTGGAAAGCAACCCC 735
Dbb 647 AATAAGTTGAGACATACAG-----GCTCGAAAGTTTATATAACCTTCGAGCCTTGTG 702
QY 736 CTGATTAACAAGGTCAGGCTACATATGCTCTCTCTCTCTGCTGAAATCCGAAAGCTAA 795
Dbb 703 GTCCCACTGATAAGTTCGTCACAGTTTGTGTGACCCGCACTAATATCAATCTACTAA 762
QY 796 AGAACAAGTACTGAATCTCAGAGGATCGGAACCGACATACGTGTAAACAGCTTCACAA 855
Dbb 763 AGAAAAGGCTTTAA-----CCCAAGTGCCTAACTTGGAGTACATGT--CATCTTTTACGG 816
QY 856 TGACGTGTGGATACGATGACATGATGCTCAAAATCAAAAGATGACGTCGTATCAGAGG 915
Dbb 817 TAATCTTGTGTTATATAGGTTGATAGGTTGATAGCGGAATCA-----CTCGTAAAAATAG 867
QY 916 AATCATCGAACGACGAAATAGCTCGAGTACTTCTAGTTTACAGCGGATTTGCCAGGAC 975
Dbb 868 GAGAAAGAAAGGCGAAGCAGGCTAGAACAGTTTATATCAATCACCATTGATGTCGATCTC 927
QY 976 TTTCTAGCGCCCGTGTCCGCTAACTACTTTGGCAACTGTCTTGGCTGATCGGTTGCCAA 1035
Dbb 928 GTCTTGATCCACCAATTTCCACAGCCTACTTTTGGTAACCTGTGGTGACCATGTGTCCTCG 987
QY 1036 AAGCAACATGAAGATAGTTAGTTGGGATATAAGGGCTTCTTGTGCGATTGCGAGTTATG 1095
Dbb 988 CTTAAAAAATGCTGTTTGTACTAGCGAAATAGGGTATGACCTTGTGTCTAAAGTAATTG 1047
QY 1096 GAGAAGCCATTGAAAAGAGGTTTCACAACGAAAGCGGTTCTTCGAGATGCAAAAACCTT 1155
Dbb 1048 GAGAGCTATATGCAAAATGATATATAATAAGCAGGAACTTCTGAAAGATGCCCGAGAT 1107
QY 1156 GGTATCGGAATCTAATGGAATCCCTTTCAAAAAGATTTCTCGGGATTAACCGATCGCCTA 1215
Dbb 1108 GG---CATGAACCTTTTCATGATCCCGGCTAGGAAGATT---GGTGTGCTGGTACACCTA 1161

QY 1216 AGTTTCGATTGCTATGTTAGATTTTGGATGGGAAAGCCTGCAAAATTTTGACATTACCT 1275
Dbb 1162 AGCTCAACTTGTACGACTTGTGATTTTGGTGGGGAAGCCGATAAAGTATGAGACTGTTT 1321
QY 1276 CTGTTGATTATGCAGAAATTTGATTTATGTTTCAGTCCAGGGATTTTGA AAAAGGTGTGG 1335
Dbb 1222 CAATAGACTATAATACGTCGATTTCTATTAATGCAAGCAAAACATCAGCACAAAGATCTTG 1381
QY 1336 AGATTGGAGTATCATTTGGCTAAGATTCATATGATGATTCATTTGCAAAAATCTTTGAAGAAG 1395
Dbb 1282 AATTGGATTGACTCTACCGAGTATGCAATGAGGCGGTTTCTAGCATCTTTGATGAAG 1341
QY 1396 GCTT 1399
Dbb 1342 GATT 1345
RESULT 4
LOCUS E12754 1622 bp DNA PAT 24-JUN-1998
DEFINITION Gentianatriflora mRNA for acyltransferase, complete cds.
ACCESSION E12754
VERSION E12754.1 GI:3251586
KEYWORDS JP 1997070290-A/2.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1622)
AUTHORS Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y.,
Yonekura, K., Mizutani, M. and Kusumi, T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 2 18-MAR-1997;
SUNTORY LTD
COMMENT OS Gentianatriflora
PN JP 1997070290-A/2
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PI
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:R65),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1622
FT /variety='japonica' /organism='Gentiana triflora' FT
FT /tissue_type='petal'
FT /clone='pgAT106'
FT CDS 35..1474
FT /product='acyltransferase'.
FEATURES
source
I..1622
/organism='unidentified'
/db_xref='taxon:3264'
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN
Query Match 11.1%; Score 189.2; DB 10; Length 1622;
Best Local Similarity 52.2%; Pred. No. 1e-37;
Matches 728; Conservative 0; Mismatches 618; Indels 48; Gaps 12;
QY 25 TGAAGGTTCTTGA AAAATGCCAAGTTTACACCACCATCTGCACACACAGATGTCAGTTAT 84
Dbb 57 TCAAGGTTCTTGA AAAATGCCGTTTTCGCCGCCACCCAGGAC---GCCGTCGCCGAGTTTA 113
QY 85 CGTACCGGTTAACATTTCTTCGATATCCCTCGTTGTCACCTTGAATAAGATGCGATCCCTTC 144

Db	114	CAGTCCCACATGTCGGTTTTTCGACATCGCGATGGTGATCTCTGATGCGAGAACCATCTGC	173
Qy	145	TGTTTTACGACTTTCCGTAACCCAGAACACATTTCTTGACACATGTTATCCCTAATCTTA	204
Db	174	ATTCTACAGATTCGCCCATCTTGTGCCAA---CTCTAAATTTATCATTTTCATCCATTA	230
Qy	205	AGGCCTCTTTGTCCTCTACTCTAAACACTAGTTCCTCTAGCGGAAATTTGTTGATGC	264
Db	231	AATCGTCCCTTTCCCTTGTTCTCAAAACATTTCTTCCGTTAGCCGGGAAATTTGATTTGGC	290
Qy	265	CGATCAATCGGGCGA---AATGCCGAAGTTTCAGTACTCCCGTGATGAGGGCCACTCGA	321
Db	291	CGGTAGATTCTCCGATAGAATCCGCGAGTTGCGTTAC-----AAGAAGGGGACTCCG	344
Qy	322	TAACTTTGATCGTTGCGGAGTCGACACAGATTTTGTACTCTACTTAAAGTGCATCAACTGG	381
Db	345	TTTCTTTAACATTTGCAGATCGAGATGGATTTTGATTTATCTCGCGGAGATCATCAGA	404
Qy	382	TAGATTCCAAATTTGCGATGCGCCTTTTTTATGTTATGCCACGCGGTTATAGGACCATGC	441
Db	405	GGGATTCTTATAAATTCACGATTTGATTCGCGACTGCCAGACCCGATTTCTAACCC---T	461
Qy	442	AGACTATAAAGTATCCGCGCTGTAAGCGTGCAAGTAACCGTTTTTTCCTAACCCTGGCA	501
Db	462	CCGGCAGCAAGTATTACCACATTTTGTCTTACAGGTGACGGTGTCTCCAAACCCGGA	521
Qy	502	TAGCGTGGCTCTGACGGCACATCTCAATTGCAGATGCTTAAAGTTTTGTATGTTCA	561
Db	522	TATGCATTTGGACGCAATCTTCAATCAAGTCTTGGTGATGCCAGTCTCTTCGATTTTA	581
Qy	562	TCAATGCTTGGGCTATATTAACAAT---TTGGGAAAGACGCGGACTTGTGTGCCGGA	618
Db	582	ATAAATTTGGGTTTGGTTGACAAATCCAATGGAGATTCATTAAAGTTCTCTCCACTTT	641
Qy	619	ATCTTTTCATCTTCGATAGATGATTAATCAAGATCTGTATGGCCCTAGAGGAACAT	678
Db	642	CTTCTCTACCTATGTACGACAGATCTGTGGTGCAAGATCCATTTTCATATTCGTGCAAAA	701
Qy	679	TTTGGAAACGAATGCAAGATG---TCTTTGAATGTTCTCTAGATTTGGAGCAAAACCCC	735
Db	702	TCTACAATGAAGAAACGTGCTCAATCTCAGGGCACACCTACTGTTCTTAATCCAGAA	761
Qy	736	CTCGATTCACAAAGGTACGAGCTACATATGCTCTCCCTTGCTGAAATCCAGAAGTAA	795
Db	762	TTTCTAAAGATGAAGTTTCGAGCCACTTCATCCTACACCTATTGATATCATGAAGTCA	821
Qy	796	AG-AACAAGTACTGAATCTCAGAGATCCGAACCCGACAATA-----CGGTAAACGA	846
Db	822	AGAAATTCATTTTCGTCAAAAAATTCGCAACTTAACCGGTAGTAGTAATTTAATCTGTCAA	881
Qy	847	CGTTCACAATGACGTCTGGATACGATATGGACATGATGTCATCAATCAAAAGATCACCTCG	906
Db	882	CTTTACGGTGACATCTGCACGTATCTGGACATGTTGTGAAATTCATTAGACACCGCTCG	941
Qy	907	TATCAGAGGAATCATCGAACGACGAAAAATGAGCTCGAGTACTTCAGTTTTCACGCGGATT	966
Db	942	TAAGAGAA---GGTGGAGAGGATAACATGCAAGCAAACTTATGTCTTTTCATCAACT	998
Qy	967	GCCGAGACTTTGACGCCGCCGTGTCGCCCTTAATCTACTTTGGCAACTGTCTTCGCTCAT	1026
Db	999	GCCGACAACGTTTGTCTCCGCCGACCTCAAAATTAATTTGGAAATTCATAGACACCGCTT	1058
Qy	1027	GCGTTCGAAAACACACATAAAGAGTATGTTGGGGATAAAGGCGTCTTGTTCAGTTG	1086
Db	1059	GTATGTGGGATCGCATATGACAACTTGTAGGAAATGAAGGTTGTGCGTAGTCCAA	1118
Qy	1087	CAGCTATTGGAGAGCCATTGAAAAGGTTGSCAACGAAAAAGGCGTCTCTCCAGATG	1146
Db	1119	CCGCCATCGAGATGCTATCCATACAGGTTACATGACTACGAAGGAATTTCTGAGAGAG	1178
Qy	1147	CAAAAACTTGGTTTATCGGAATCTA-----ATGGAATCCCTTCAAAAAGATTTCTCG	1197

Db	1179	ATTGGATATACCGCCCGGATCAACATCTGCGGCACCAAGTCGACGCTCATTTATGTCG	1238
Qy	1198	GGATTACCGGATCGCCTAAGTTCGATTCGTTAGTATTTTGGATGGGGAAGCGTG	1257
Db	1239	TTGGATCGGCACACGCAATTCGATGATTTTGTATGATGAGATTTTGGTGGGAAAGCTTG	1298
Qy	1258	CAAAATTTGACA---TTACCTCTCTGTTGATATGACAGAAATGATTTATGTCAGTCCGCA	1314
Db	1299	AAAAGCATGAATCTGTTTCAACTAATCCTTCGGCAACACAAATTTTGTATCTCTCGGTCCA	1358
Qy	1315	GGGATTTTGGAAAAGGTCTGGAGATTCATTTGGCTTAAGCATCATATGATGATGCAT	1374
Db	1359	GAAATTTAAAGGAGCAGCTTGAGCTTGGCATTTCTTGGCTAAGAATAGGATGGACCAT	1418
Qy	1375	TGCAAAAATCTTT	1388
Db	1419	TGCGACCATTTT	1432
RESULT	5		
LOCUS	AB026494	1622 bp	PLN
DEFINITION	Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete cds.		
ACCESSION	AB026494		
VERSION	AB026494.1	GI:7415596	
KEYWORDS	acyltransferase homolog.		
SOURCE	Gentiana triflora cDNA to mRNA.		
ORGANISM	Gentiana triflora		
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.		
AUTHORS	1 (sites)		
TITLE	Ionekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M., Fujiwara,H., Fukui,Y., Toshihiko,A., Yamaguchi,M. and Kusumi,T. Molecular cloning and biochemical characterization of hydroxycinnamoyl-CoA:anthocyanin 3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla frutescens and diverse plant acyltransferase homologs unpublished (1999).		
JOURNAL	2 (bases 1 to 1622)		
REFERENCE	Tanaka,Y. and Yonekura-Sakakibara,K.		
AUTHORS	Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.		
JOURNAL	Yoshikazu Tanaka, Suntory Ltd., Institute for Fundamental Research; Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan (E-mail:Yoshikazu.Tanaka@suntory.co.jp, Tel:81-75-962-8807, Fax:81-75-962-8262)		
FEATURES	Location/Qualifiers		
source	1..1622		
gene	/organism="Gentiana triflora"		
CDS	/db_xref="taxon:55190"		
	35..1474		
	/gene="GAT106"		
	35..1474		
	/codon_start=1		
	/product="acyltransferase homolog"		
	/protein_id="BA093452.1"		
	/db_xref="GI:7415597"		
	/translation="MAGNSDDIKVLEKCRVAPPDDVAEPTVLPSTFDMRWLISDAEH HLHFRRPCPKSKFIISKLSLVLKHFLPAGNLITWVDSDDRMPELRYKKG SVLTTAESMEDYLAGDHQDSYKFNLDIPOLPEPTVSGDEVLPFLAQVTFSN TGICIGNLHQVLGDASSFHFKNLWLVNDKSGSLKFLPLSSLPWDRSVVDDPFI IRKRIYNERKLKSQGTPTVLNPAISKDEVRA7F ILHPIDIMKKKFISSKNRLTGS SNYNLSTFTVTSALWTCLSLDTVVREKVEDKHAANLCARCFORFAPP1PQNY FGNCTVPCWGS7HEQLVNEGNSVAATGADIAHKRLDIYEGILRGDWISPPRSTSA APSTLIYVVGSAQRNVHDFDAGWGKLEKHESVSTNPNSATILITISRRRFRKGALE GISPKNRMDAFATITITFINSLVHRSPL"		
BASE COUNT	478 a	345 c	477 t
ORIGIN			

Query Match 11.1%; Score 189.2; DB 12; Length 1622;
Best Local Similarity 52.2%; Pred. No. 1e-37;
Matches 728; Conservative 0; Mismatches 618; Indels 48; Gaps 12;

QY	25	TGAAGGTTCTTGAANAATGCCAAGTTACACCACCACCTCTGACACAACAGATGTCGAGTTAT	84
DB	57	TCAAGATTCTTGAGAAATCCGTTGTCGCCACACCGGAC---GCCGCGCGGAGTTTA	113
QY	85	CGCTACCGGTACATCTTTCGATATCCCTCGTGTGCACTTGAATAGATGCACTCCCTTC	144
DB	114	CAGTCCCACTGCTGTTTTTCGACATCGGATGGTTGATCTCTGATGAGAACACCATCTGC	173
QY	145	TGTTTTFACAGCTTTCGATACCCCAAGAACACATTTCTTGGACACTGTTATCCCTTAATCTTA	204
DB	174	ATTCTACAGATTCCGCCACTCTTGTCCCAA---CTCTAAATTTATCATTTTCATCCATTA	230
QY	205	AGGCTCTTTGCTCTCACTCTAAACACACTAGTTCGCTTACCGGAAATTTGTTGATGC	264
DB	231	AATCGTCCCTTTCCCTTGTCTCAACACACTTCTTCGCTAGCCGGAATTTGATTTGGC	290
QY	265	CGATCAAAATCGGCGCA---AATGCCGAAGTTTCAGTACTCCCGTGATAGGGCGACTCGA	321
DB	291	CGGTAGATTCTCCGATAGATGCCGGAGTTGGTTAC-----AAGAAGGGGACTCGG	344
QY	322	TAACTTTGATCGTTGGGAGTCTGACAGGATTTTGACTACCTTAAAGGTCAATCAACTGG	381
DB	345	TTTCTTTTAACAATTGCAGAAATCGAGCATGGATTTTGCATATCTCGCGGAGATCATCAGA	404
QY	382	TAGATTCCAATGATTTGATAGGCTTTTTTATGTTATGTCACGCGGTTTATAAGGACCATGC	441
DB	405	GGGATTTCTTATAAATTAACAGATTTGATTTCCGCGAGCTGCCAGAACCGATTGTAACC---	461
QY	442	AGACTATATAAGTGATCCCGCTGCTAGCGTGCAAGTAAACCGTTTCTCAACCGTGGCA	501
DB	462	CGGCGCAAGATATTACCACTTTTTCGCTTTACAGTGACGGTGTCTCCAAACCCGGTA	521
QY	502	TAGCCGTGCTCTGACGGCACATCAATTTGCAGATGCTAAAAGTTTGTAAATGTCA	561
DB	522	TATGCAATTTGGACGCAATCTTCATCAAGTTCTTGGTGATGCCAGTTCTTTCTGCAATTTTA	581
QY	562	TCAATGCTTTGGGCTATATTAAACAAT---TTGGGAAAGACGGGACTTGTGTCCGCGA	618
DB	582	ATAAATATTGGGTTTGGTTGACAAATCCAAATCCAAATGGAGATTCATTTAAAGTTCCCTT	641
QY	619	ATCTTCTTCCATCTTTCGATAGATCGATTAATCAAGATCTGTATGCCCTAGAGGAAACAT	678
DB	642	CTTCTCTACTATGTACGACAGATCTGTGGTGCAAGATCCATTTTCATATTCGTGCAAAA	701
QY	679	TTTGGAAACAAATGCAAGATGT---TCTTGAATGTTCTCTAGATTTTGGAAAGCAAAACCC	735
DB	702	TCTACAATGAAGAAAATGCTCAAAATCTCAGGGCACACCTACTGTTCTATAATCCAGCAA	761
QY	736	CTCGATTCAACAAGGTACGAGCTACATATGCTCTCTCCCTTGGTGAATCCAGAAGCTAA	795
DB	762	TTTCTAAAGATGAAGTTCGAGCACCTTCACTCTACACCCCTATTGATATCATGAAGCTCA	821
QY	796	AG-AACAAGTACTGAATCTCGAGAGATCCGAACCCACAATA-----CGGTAAACGA	846
DB	822	AGAAATTCATTTGCTCAAAAATTCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA	881
QY	847	CGTTCAATAGCGTGTGATAGTATGGACATGATGGTCAATTAATCAAAAGATGAGCTCG	906
DB	882	CTTTACGGGTGACATCTGCACATGATCTGGACATGCTTGTGCAATCATTAGACACCGCTG	941
QY	907	TATCAGAGGATCATCGAAGGCAAAATGAGCTCGAGTACTTCAGTTTACAGCGGAT	966
DB	942	TAAGAGAGAA---GGTGGAGAGGATTAACATGCACAAACTTATGTGCTTTCATCAACT	998
QY	967	GGCGAGGACTTCTGACGCCCTGTCGCGCTAACTACTTTGCAACTGCTTTCGCTGCAT	1026
DB	999	GCGCAACAGTTTGTCTCCGCCCTACCTCAAAATTTACTTTGGAATTTGCATAGTGCCTT	1058
QY	1027	GCCTTGCAAAAGCAACACATAAAGAGTTAGTTGGGGATAAAGGCTTCTTGTGCAAGTTG	1086

DB	1059	GTATGTTGGGATGCACTCATGAGCAACTTGTAGGAAATGAAGGTTGTCGGTAGCTGCAA	1118
QY	1087	CAGCTATTGAGAGCAACATTGAAAGAGGTTGCACACGAAAGGCGTTCTTTCGAGATG	1146
DB	1119	CGCCCATCGAGATGCTATCCATAAGAGGTTTACATGACTACGAGGAATTTCTGAGAGGAG	1178
QY	1147	CAAAAATTTGGTTATCGGAATCTA-----ATGGAATCCCTTCAAAAAGATTTCTCG	1197
DB	1179	ATTGGATATCGCGCCCGCATCAACATCTCGCGCACCAAGTCCGACGCTCATTTATGTCG	1238
QY	1198	GGATTACCGGATCGCCTAAGTTCGATTCGTATGTTGATGTTTGGATGGGAAAGCCTG	1257
DB	1239	TTGGATCCGCACACGCAATGTCATGATTTTGTATGATGATTTTGGTTGGGAAAGCTTG	1298
QY	1258	CAAAATTTGACA---TTACCTCTGTTGATATGCAGAAATTTGATTTGATTCAGTCCCA	1314
DB	1299	AAAGCATCAATCTGTTTCAACTTAATCTTCGGCAACACATAATTTTGATCTCTCGGTCCA	1358
QY	1315	GGGATTTTCAAAAAGGTGTGGAGATTGGAGTATCATTTGCTTAAGATTCATATGGATGCAT	1374
DB	1359	GAAGATTTAAAGAGGACACTTGAGCTTGGCATTTCTTGGCTAAGATAGGATGCGGCAT	1418
QY	1375	TTGCAAAAATCTTTT	1388
DB	1419	TTGCCACCATTTT	1432
RESULT	6		
E12756			
LOCUS	E12756	1479 bp	DNA
DEFINITION	Perilla ocimoides mRNA for acyltransferase, partial cds.		24-JUN-1998
ACCESSION	E12756		
VERSION	E12756.1	GI:3251588	
KEYWORDS	JP 1997070290-A/4.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1479)		
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.		
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY		
JOURNAL	Patent: JP 1997070290-A 4 18-MAR-1997;		
COMMENT	SUNTORY LTD		
	OS Perilla ocimoides		
	PN JP 1997070290-A/4		
	PD 18-MAR-1997		
	PF 30-JAN-1996 JP 1996046534		
	PI 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915		
	ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO,		
	PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAHIKI		
	PI C12N15/09, A01H1/00, C07H21/04, C07K14/42, C12N9/10, (C12N9/10, PC		
	C12R1:865),		
	PC (C12N9/10, C12R1:19);		
	CC strandedness: Double;		
	CC topology: Linear;		
	CH key	Location/Qualifiers	
	FT source	1. 1479	
	FT /organism=	'Perilla ocimoides'	FT
	FT /tissue_type=	'leaves'	
	FT /clone=	'psAT208'	
	FT CDS	1..1343	
	FT /product=	'acyltransferase'	
FEATURES	Location/Qualifiers		
source	1..1479		
	/organism=	'unidentified'	
	/db_xref=	'taxon:32644'	
BASE COUNT	420 a	316 c	331 g
ORIGIN	411 t	1 others	

BASE COUNT 419 a 315 c 331 g 411 t
ORIGIN

Query Match 10.1%; Score 171.4; DB 12; Length 1476;
Best Local Similarity 49.9%; Pred. No. 3.6e-33;
Matches 707; Conservative 0; Mismatches 661; Indels 49; Gaps 9;

Qy 78 GAGTTATCGCTACCGGTAAACATCTCTCCATATCCCTGGTTCGACCTTCAATAAGATCGAG 137
Db 48 GAGCAATCGGTCCGCTCACATCTCTGACATGAGCTGGCTGATCTTTTCAATCCCATGCTT 107
Qy 138 TCCCTTCTGTTTACGACTTTCCGTACCAAGAACACATCTTCTTGGACACTGTTATCCCT 197
Db 108 CAGCTCCTCTCTACGAATTCCTTGTTCGAAGCAACATTTTTCAGATCCATCGTTCCA 167
Qy 198 AATCTTAAGCGCTCTTGTCTCTACTCTAAACACTACGTTCCGCTTAGCGGAATTTG 257
Db 168 AAACCTCAACAACATCTCTCTAATAACTCTCATACACTCTTCCCTCTCTCATGCAATTTA 227
Qy 258 TTGATGCGCGATCAAAATCGGCGAAATGCCGAAGTTTCAGTACTCCCGTGATGAGGCGAC 317
Db 228 ATCTACCTTTCATCCCCGGAGAAATGCCGGAGTTTCGGTA-----TCTATCCGGGAC 281
Qy 318 TCGATAACTTTGATCGTTCCGGAGTCTGACCAGGATTTTGTACTACTTTAAAGGTTCATCAA 377
Db 282 TCGGTTCTTTCACCATCGCAGAACTAGCGACGACTTCGATGATCTCGTCGGAATCGT 341
Qy 378 CTGGTAGATTCGAATGATTTGATGGCTCTTTTATGTTATGATGCCACGGGTTATAAGGACC 437
Db 342 CCAGAATCTCCCGTTAGGCTCTACAACTTTGTCTCTAAATTTGCCGCCATTTGTCGAAGAA 401
Qy 438 ATGCAAGACTATAAAGTATCCCGTCTGAGCGGCAAGTAAACGTTTTCCTTAACCGT 497
Db 402 ---TCCGATAGAAAACCTCTCTCAAGATTTTCGGCGGTGAGGTGACTCTTTCCCGAGCCGA 458
Qy 498 GGCATAGCGGTGGCTCTGACGGCACATCATTTCAATTTGCAGATGCTTAAAGTTTGTAAATG 557
Db 459 GCGTCGCTATTGGAATAGCAACGATCACACCGTTAGCGAGCGCCGCTCGTTCTCGCG 518
Qy 558 TTCATCAATGCTTTGGGCTATATTAAACAAA-----TTTGGGAAGACGGGACTGTTGTGCC 614
Db 519 TTATATAAGCGCTTGCTCTCAATGAGCAACACATTTGAAATGAAGATGAAGATGAAGAA 578
Qy 615 GCGAATCTTCTCCATCTTCGATAGATCGATAAATCAAGATCTGTATGGCTAGAGGAA 674
Db 579 TTATAATCTTCCAGATTTTCCGATAGATCCGTCATAAATATCCGACGAAATTTGACTCC 638
Qy 675 ACATTTTGAACGAATGCAAGATGTTCTTTGAAATGTTCTCTAGATTTTGGAAAGCAAAACC 734
Db 639 ATTTATTGGAGAACCGGCTAAATTTCT-----TTGCAATCTCGTCATCCCTCA 689
Qy 735 CTTCAATTCACAAGGTAGGACTACATGATGCTCTCCCTTCTGCTGTAATCAGAGCTA 794
Db 690 TTACCGACGCGGCGGATTTGGAACACCGTTCTGTTTTCACCCCAATCCAAAATTAAGAATG 749
Qy 795 AAGACAAAGTACTGAATCTCAGAGATCCGAACCGACAAATACGTTACGACGTTCA 854
Db 750 AAG-----GGTTGGATTGAGTCAGAGTTCCAAAGTTTATGTCATCTCTCTTTTGT 803
Qy 855 ATCAGCTGTGGATAGGTATGGACATGCCATGTTCAATCAAAAGATGACGTCGTATCAGAG 914
Db 804 GCGATTGACGCTTATATGTGGCTGGCATAACGAATCATTCACAGCAGTAA-----857
Qy 915 GAATCATCGAAGACGCAAAATGAGCTCGAGTACTTCAGTTTACAGCGGATTTGCGGAGGA 974
Db 858 -----GACCAAGACAAACGAGGATGATTTTCTTGTATTCGCTCGATCTAAGGCCA 908
Qy 975 CTTCTGACGCCCGGTGTCGGCTTAAGTCTTGGCACTGCTTTCGCTGATCGCTTCA 1034
Db 909 CGATTAGATCCCGCGGTTCTCTGAAATTAATCTTCGGGAAGTCTTATCGTACGCGCTGCCG 968
Qy 1035 AAAGCAACACATAAAGAGTTAGTTGGGGATAAAGGGCTTCTTGTTCAGCTTTCAGCTATT 1094

Db 969 AGAATGCGCGCGAGAGCTGTGGGAGAGAAAGGGTGTCTTCTGCAGCTGAGTAATC 1028
Qy 1095 GGAGAAGCCATTGAAAAGAGGTTGCACAAAGAAAAGCGGTCTTTCGAGATGCCAAAAC 1154
Db 1029 CGCGCGGAGATAAAAAAGGAT---CAACGACAAGAGAATATTAGAAACGGTGAGAAA 1085
Qy 1155 TGGTTATCGGAATCTAATGGAATCCCTTCAAAAAGATTTCTCGGATTAACCGATCGGCT 1214
Db 1086 TGGTCCGCGAGATTCGTAAACGTTGCAGAAATCATATTTTTCGGTGGCAGATCGAGC 1145
Qy 1215 AAGTTTCGATTCGTATGTTGATTTGGATGGGAAAAGCCCTGCAAAATTTTGACATFACC 1274
Db 1146 AAGCTAGATCTTTACGTCGCAGATTTTGGATGGGGAAGCGGAGAAAGCAAGAAATATTG 1205
Qy 1275 TCTGTTGATTAATGACGAATTTG---ATTTATGTGATTCAGTCCAGGATTTTCAAAAAGT 1331
Db 1206 TCGATTGATGGGGAATAATGCAATGACACTTTCTAAGCCAGGATTTCCAGAGGAGA 1265
Qy 1332 GTGGAGATTGGAGTATCATTTCTAAGATTCATATGATGATGCAATTTGCCAAAATCTTTGAA 1391
Db 1266 TTGGAGTTTCTGCTCTTTTSCCTAAGGACAAAATGGATGCTTTTCTGCTTATTTTTCA 1325
Qy 1392 GAAGCTTTTGTCTTCTGCTATGATCTCTTTAATAGAACCATATTTTGTGCAATAAAGTA 1451
Db 1326 CTGGCAATTAATGTTAATAATGTAT-GTAATTAACCTAATATTGTTGCTGTAACAATTA 1384
Qy 1452 CCAAGTCTTTTGTAGTAACACTACACCAAAACCTTACTTT 1488
Db 1385 ATTAAGTGTGAGTAACCTGAAGAATAATACCTTATTAT 1421

RESULT 8

AC003027 119914 bp DNA PLN 30-JAN-1999
LOCUS Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,
DEFINITION complete sequence.
AC003027
AC003027.1 GI:4079614
VERSION HTG.
KEYWORDS
SOURCE Arabidopsis thaliana
ORGANISM
REFERENCE 1 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altai,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,
Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vysotskaia,V.,
Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
REFERENCE 3 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altai,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1998) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
REFERENCE 4 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,

Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P., Gonzalez,A., Kremetska,I., Kim,C., Lenz,C., Li,J., Liu,S., Luross,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V., Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
 Direct Submission
 Submitted (30-JAN-1999) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA

TITLE JOURNAL

COMMENT

On Dec 30, 1998 this sequence version replaced gi:2734094.
 Bases 1-9262 of clone F21M11 overlap with bases 68998-78259 of 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7 (AC002560).

e-mail for correspondence: arabidopsis@sequence.stanford.edu
 Genes with similarity to proteins in the databases are described as 'putative', '-like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The software programs used to predict genes include: Grail (informatics.group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge, <http://gnomic.stanford.edu/~chris/GENSCANW.html>), Fexa (V.Solovyev & A.Salamov, Sanger Centre, <http://genomic.sanger.ac.uk/>), and NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/NetPlantGene.html>).

FEATURES

source

1..119914
 /organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="I"
 /clone="F21M11"

misc_feature

1..7000

/note="Proteins in this region are annotated in the F21M11 entry, AC002411."

misc_feature

1..9262

/note="overlap with bases 68998-78259 of 'IGF' BAC clone F21M11."

gene

7684..10913

/gene="F21M11.1"

join(7684..7866,7965..8027,8122..8160,8250..8381,

8458..8728,8875..9893,10003..10256,10344..10410,

10501..10551,10685..10781,10885..10913)

/gene="F21M11.1"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010665.1"

/db_xref="GI:4204284"

/translation="MAPNLRIRKACDAMKLLGISETKTRAFRLKLLKTYENNWFTEE
 DAYKVLDAIFDADAQSTKKNKEKKKKKKKSVATSRGRKKAPELVQDEE
 DMDDEFFPLKRLRRGRASSSSSSSYNNEDLKTQPEEDEDGVTPLPKRY
 VRNGEGLATVYNNASPSSSRLSMEPEVPMVLLPAHPMETKVSASALVLLD
 EPNIHKPVISDTGNCAPSMLKGNHIVQEMWETKDLNDITAMDVSPSAIGES
 SEHKVAASVELASSTSGEAKICLSPATGETTNLHPSMEDLRAMEECLKYSKI
 VHPFESVLGMCSCYCIDLAKNSTQLLETETCDMSKAGDESAVGTSMPLVYP
 ECISGDGKAIKMDKIDTAGENVEIPWNEINEKVPFRYPHPSFVQDAVIFDS
 LSSFSQSCSTSCIEDCLASEMNCNCAICVDNGEAVTLGLLKEFLEARISEARDQ
 RKQVLRFCCEPLERAKKVEILEPCGHLKRGAIKCEWPKCGCTKRCGRNVYVQGMHN
 KLVQVFPNGKGLRLEKLPKGAFCICEYIGELTIPELYQSFEDKPLPVILDAH
 WGEERLEGKALCLDGMFVGNISRLNHRDLNLTETPVQVETPDQHYHAFETT
 RQTEAMEELAWDYGIDFNNDNSLMKPFDCLCGSRFCRNKRSTKTQMLNKA"
 join(13033..13488,13717..13926,14139..14514)

mRNA

13033..14514

/gene="F21M11.2"

join(13061..13488,13717..13926,14139..14316)

/gene="F21M11.2"

/note="Similar to acid phosphatase; Location of ESTs

110C2T7, gb|T42036, and 110C2XP, gb|A1100245"

/codon_start=1

/protein_id="A010666.1"

/db_xref="GI:4204285"

/translation="MDRTWFLSLTIAASLLVGVVVSAGDWNILLNQLRGLGSSSSQNGIVS
 KGITLKGKCESARINVEVHNIRKEDVVPQECVSHIKDWTSSQKDDVARTVDEVI
 LHFSGCCSKCDGMDAMTIFDIDTLLSTIPYHKKNGFGGKGLNSTKQFADWIKQK
 APAPVHKLYHDIRERGIKFLISSRKVELRSATVDNLIOAGYVGNLSNLIYYVA"
 OREVKQYKSEKKRWLMSLGYRVWGVMDQSSPAGCPLPRRTFKLNSLIYYVA"
 complement(join(14004..14312,14565..14641,14787..15831,
 15906..16006,16327..16362))
 /gene="F21M11.3"

mRNA

gene

complement(14004..16362)

/gene="F21M11.3"

complement(join(14567..14641,14787..15831,15906..16006,

16327..16362))

/gene="F21M11.3"

/note="Unknown protein; Location of ESTs 40C3TT,

gb|AA728590 and40C3T7, gb|T04573"

/codon_start=1

/protein_id="A010667.1"

/db_xref="GI:4204286"

/translation="MLDQISGNHTHEKLSVETAEPHHLNDRVHIVEIPKASVIPITE
 ICDAEEKSPSTISKRVTFSKVTYEHVSEVSELSSEKNEEVESEKRSLSKSK
 TDOIIEVASNSGSGYPENHRKYNCRSDDDIEDEDFDCSDSDLEDEEYISDVGFSE
 DSLHNPTEKVYTDIGDKTEIDSKLRRSNETVRDGNHYDVGQVLNPEVNLTKWSAK
 SKGRTKQKQKRENSNFADQEEKRSSFGTDPODDITLSVKPKCRPEPKLRQKE
 LAVDASLWSLSTSESSECSASMTLTPEKLSKTSYKSKPLRINHDDRPVLCALTL
 EDIKQFSATSTPRKSPSKSPDETPIITGTGGYWGNSKKAIDCGSASFSGIPNTSSKY
 REDKSYNWHSTPEARLEKALNNIDK"
 complement(16835..20238)

gene

complement(16835..20238)

/gene="F21M11.4"

complement(join(16835..17185,17274..17392,17491..17656,

17795..17885,17982..18079,18175..18361,18504..18604,

18705..19049,19134..19349,19439..19791,19862..19967,

20188..20238))

/gene="F21M11.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010669.1"

/db_xref="GI:4204288"

/translation="MGSSPPEARARAQVPSMILFLEIICTVHVYNNRRKLNDRVLISA
 NLNIPKRVTRDLRMEGSEVISYKSLNAAVSVHLPOLOIPNDCNYKNALNNSPKH
 GESEDEMTDKOVSKRGSGTSSRDGSP.LPTSEESDRPKHODWTEKOLSDHLLVE
 FSEYDAANHTPSTYEQAKNVDRDITASEQPSNAARKICGDSFQIESPNPQTQDP
 TLLRLMESRSDPTDYVKAQNHQWYF.SFWIYISFVLCRLSNLKKCVIFVCQQLPK
 SHTEQSKRRDITASDAMENHLKVPKRENNLMQKSDIDCNGKCSANSDDQISEKIS
 KALEQTSNITTCGFCQCSARVSATGEMLHYSRGREVDGDDIERSVNIHVHSACIEWA
 POVYEGDVTVKLKAELAGMKIKCTKSLKGAALGCFVKSCRRSYHVCAREISCR
 WYDEFLLCPAHSSVYKFEKNSKSHVRAEPLKINPAELCSLEQTPAFTKELVLCG
 SALSXDKKLMESLAVENATISRYNPSVTHVIASTDKEGACTRLKVLMLGTLNKKW
 IINAAWKASLKASQPVDEEPEFIQIDGCGQCPKTAARLRAETNPKLFEGLKFFTF
 GDYKGYKEDLQNLVKVAGGTILNTEDELGAESSNNVNDORSSIVYVNIIDPPHGCAL
 GEVYTIWORAADAELASQTSRLVGHWTWVLESIAGYKLHPVIG"
 20752..24647

gene

complement(20752..24647)

/gene="F21M11.5"

join(20752..20994,21362..21497,21596..21740,21825..21994,

22102..22178,22542..22758,22920..23064,23343..23442,

23599..23693,24021..24072,24227..24298,24441..24647)

/gene="F21M11.5"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010668.1"

/db_xref="GI:4204287"

/translation="MGANSKSVTASFTVIAVFFLCGGRTAVDETEFHGDYSKLSGI
 IIPFASTOLRAMSILDCPYTPLDFNPLDLVLDITKLLSANCWFKCVMYLDYNGTD
 HPCKSRDGLSAITELDPGYITGPLSTWMLKWCVEFGTEANAIVAVPDKPRLS
 PKLEERDLYFHKLKLTFTETAKLRGSPSIVFAHSMGNVRYFLEWLRLAPKHYL
 KWLDOHIIHAYFANGAPLLGSAVEAKSTLSQVTFGLPVSEGTARLLNSPSSILMLPF
 SKNKGDNFTWFHFGSAAKOKRVTHCDEEYQSKYSGWPTNIIIEIPSTSARELA
 DGTFLKAIEDYDDPSKRLHLQKKYVFFVIRNIAHRSLAGLLEHDDPVPNPLTPW
 ERPKINFCIYGAHLKEVGYFAPSGRPYDNWITIIIEYTESGLSVRSQTVVDG
 NAGPIQDETVPYHSLSWCKNMLGPKVNITMAPIILIGIKQOPEHGGSDVHVELNVD
 HEHGSDIANMTKAPRVKVIITFVDESDSEIPGKRTAVWELDKSGY"
 27777..28734

mRNA

27777..28734

/gene="F21M11.6"

27777..28734

/gene="F21M11.6"
 28007. .28465
 /gene="F21M11.6"
 /note="Unknown protein; Location of ESTs 20312477,
 gpH76794 and 2031244P, gb|AA605510"
 /codon_start=1
 /protein_id="AAD10670.1"
 /db_xref="GI:4204289"
 /translation="MNTKTRLPRLVLTADKRERDAFISSVTDNPPPIAKFSPPP
 KLVPPVPISSKSTAAEPIGSLAGLYLSHEYLTQGLFGEQWNAQAQAES
 KIKPSHTVEAECEPRKRRREVANLURSDGQPLGVNPAQLARFLK"
 complement(29264. .32033)
 /gene="F21M11.7"
 complement(join(29264. .31015,31312. .31414,31484. .31587,
 31782. .32033))

Query Match 6.7%; Score 114; DB 12; Length 119914;
 Best Local Similarity 46.2%; Pred. No. 2.8e-18;
 Matches 612; Conservative 0; Mismatches 685; Indels 27; Gaps 6;

QY 85 CGCTACCGGTAACATCTTCGATATCCCTGGTTGGCACTTGAATAGATGCAGTCCCTTC 144
 DB 46008 CTCCTCCTCCTACCTCTTCGATGCCCCCTGGCTCTCTCTCCACGCGCGATCTCTCT 46067
 QY 145 TGTCTTACGACTTTCGGTACCCCAAGAACACATTTCTTGGACACCTGTTATCCCTTAATCTTA 204
 DB 46068 TCTTCTCTCTTACCAAACTCAACTGAATCTTCTCCTCCAAGACTTTGTACCAACCTCA 46127
 QY 205 AGGCTCTTTGCTCTACCTCTAAACACTAGTTCCCGCTTACGGGAATTTGTTGATGC 264
 DB 46128 AACATCTCCTCTCCATCTCTCAACATTTCTTCCCTTACGCCGGTAACCTGATTATCC 46187
 QY 265 CGATCAAAATCGGCGAAATCGCCGAAGTTTCAGTACCTCCGCTGATAGGCGGATCGATAA 324
 DB 46188 CGCTCGTCCGAC-----CTCCATATTTACACTTAAACGATGCCAAGACTCTCTTG 46241
 QY 325 CTTTGATCGTTGGGAGT---CTGACCAAGATTTTGGACTTACCTTAAAGGTCACTCAACTGG 381
 DB 46242 TTTTACCGGTAGCAGAGTCTACTGAAACCGACTTCGACCACTGAAGTCTGATTCCCTCA 46301
 QY 382 TAGATCCCAATGATTTGCATGCGCTTTTATATGTTATGCCAGGGTTATAGGACCATGC 441
 DB 46302 AGATATCAGTGTGTGATGCGCTCTTCCCAAGTTTACCTCTCTCAGTCTCTCCGG 46361
 QY 442 AAGACTATAAAGTGATCCCGCTGCTAGCGGTGCAAGTAAACGGTTTCTTCAACCGTGGA 501
 DB 46362 AAGGAATTCAGATGCGACCAATAAATGGCTATGCAAGTCAACCTCTCCCTGGAGCTGGA 46421
 QY 502 TAGCCCTGGCTCTGACGGCACATCATTTCAATTCAGATGCTAAAAGTTTGTAAATGTCA 561
 DB 46422 TCTGTATAGCAACTCAGCTACACATGTTGTAGCAGATGGAGTCACTTTCAGTCAATTCA 46481
 QY 562 TCATCTCTGGGCTATATTAACAATTTGGGAAGACCGGGACTTGTGTCCGCGCAATC 621
 DB 46482 TGAAGTATTGGATGTCAATTGACCAATTCAGCGGTAAAGATCCCGCACGGTTCTTCTAC 46541
 QY 622 TTCTTCCATCTTTTCGATAGATGCATATCAAGATCTGTATGGCTAGAGGAACATTTT 681
 DB 46542 CCTCTCTGCTATTACAGCTGCGAAGACATGATCAAGGACCCGCGAGGTAGGCGCAG 46601
 QY 682 GGAAACGAATGCAAGATGTTCTTGAATATTTCTTCTAGATTTTGAAGCAACCCCTCGAT 741
 DB 46602 GACATTTGGAGCGGTTTTCGAGCCAAACTCTGCAAAACACAGTTTCACATGTTACGCCGTG 46661
 QY 742 TCACAAGGTAGCAGGTACATATGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 801
 DB 46662 AGAACATGCTCAGAGCCACTTTTACATTTAGCGCGAAGCAGATAGATAATCTAAAGATT 46721
 QY 802 AAGTACTGAATCTCAGAGGATCCGACCCACAAATAGTCTTAACGAGTTTCAACATCACT 861
 DB 46722 GGGTTACAGAC-----AGTCTGAAGATCAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 46775
 QY 862 GTGGATACCTATGGACATGCATGGTCAAAATCA---AAGATGACGCTCGTATCAGAGGAAT 918

DB 46776 TAGCTTTTCATTTGGGTGAAGCTTGATCAAGACACTTGTACAAAGACACTGAAACAAAGGCCA 46835
 QY 919 CATCGAACACCAAAATGAGCTCGAGTACTTCACTTCACTTACAGCGGATTCGCCAGGACTTC 978
 DB 46836 ACGAGGAAGACAAAGGATGAAGTATTTCACCTTGATGATCAATGTGGATTCGAGGAATCGCC 46895
 QY 979 T-----GACGCCCGCTGCTCCGCTAACTACTTTGGCAACTGTCTTGGCTCATGCGTTG 1032
 DB 46896 TCAAGTACACACCAACCAATACACAAACATCTTTGGCAACTGTATGGCTCTCTGATCG 46955
 QY 1033 CAAAACCAACACATAAAGAGTTAGTTGGGATAAAGGGCTTCTTGTTCAGTTTGCAGCTA 1092
 DB 46956 TATCAGCTCAAGAAACACAGATTTGCTAGGAGAAAAATGCGCTTTTGGCGCTTCAGATGCAA 47015
 QY 1093 TTGGAGACCCATTGAAAAGAGTTGCCAACGAAAAAGCGCTTCTTCAGATGCAAAAA 1152
 DB 47016 TCACAGCGAGAATCAAGATATGTTATCAAGCG---ATCTGTTGAAGACAGCACCAAGAT 47072
 QY 1153 CTTGGTTATCGGAATCTAATGAATCCCTTCAAAAAGATTTCTCGGGATTTACCGGATCGC 1212
 DB 47073 GGGACAAAGAGTACGTAATGGTCTATGCCATTTATCAACCTCAATTCGAGGAGCTC 47132
 QY 1213 CTAAGTTCGATTCGTATGCTGATGATTTTGGATGGGAAAGCGCTGCAAAATTTGACATTA 1272
 DB 47133 CGAAATTTGGGACTGTATGACATGGATTTTGGATTAGGAAAGCCCTGCAAAATGGAGATCG 47192
 QY 1273 CCTCTGTTGATTTATGCAGAAATGATTTATGTTAGTTCAGTCCAGGGATTTTGAANAAGTG 1332
 DB 47193 TGCACATAGAAACAGGTGTTCTTCGCAATCTCAGAGTCCAGAGACGGCAGCAATGGAG 47252
 QY 1333 TGGAGATTCGAGTATCATTTCCCTTAAGATTCATGATGATTCATTTGCAAAATCTTTTGAAG 1392
 DB 47253 TTCAGATTCGATAGCACTAGAGAAGAAATGATGTTGTTGACTCAATTCATCAAC 47312
 QY 1393 AAGG 1396
 DB 47313 AGGG 47316

RESULT 9
 AC002560 101176 bp DNA PLN 19-JUL-2000
 LOCUS Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome
 DEFINITION I, complete sequence.
 ACCESSION AC002560
 VERSION AC002560.2 GI:9211167
 SOURCE HTG.
 ORGANISM Arabidopsis thaliana
 thale cress.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (bases 1 to 101176)
 AUTHORS Khan,S., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Kim,C.,
 Shinn,P., Altafi,H., Bei,Q., Chin,C., Chlou,J., Choi,E., Conn,L.,
 Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B.,
 Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Liu,S., Mukharsky,N.,
 Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A.,
 Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A.,
 Theologis,A. and Ecker,J.R.
 Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome
 I
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 101176)
 AUTHORS Ecker,J.R.
 TITLE Direct Submission
 JOURNAL Submitted (25-SEP-1997) Arabidopsis thaliana Genome Center,
 Department of Biology, University of Pennsylvania, 38th Street and
 Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
 REFERENCE 3 (bases 1 to 101176)
 AUTHORS Ecker,J.R.
 TITLE Direct Submission

```

JOURNAL Submitted (13-NOV-1997) Arabidopsis thaliana Genome Center,
REFERENCE Department of Biology, University of Pennsylvania, 38th Street and
AUTHORS Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
TITLE 4 (bases 1 to 101176)
JOURNAL Ecker,J.R.
REFERENCE Direct Submission
AUTHORS Submitted (26-JAN-1998) Arabidopsis thaliana Genome Center,
TITLE Department of Biology, University of Pennsylvania, 38th Street and
JOURNAL Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE 5 (bases 1 to 101176)
AUTHORS Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C.,
Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N.,
Hoving,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
TITLE Direct Submission
JOURNAL Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center,
REFERENCE Department of Biology, University of Pennsylvania, 38th and
AUTHORS Hamilton Walk, Philadelphia, PA 19104-6018, USA
TITLE 6 (bases 1 to 101176)
JOURNAL Ecker,J.R.
REFERENCE Direct Submission
AUTHORS Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center,
TITLE Department of Biology, University of Pennsylvania, 38th Street and
JOURNAL Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE 7 (bases 1 to 101176)
AUTHORS Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C.,
Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N.,
Hoving,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
TITLE Direct Submission
JOURNAL Submitted (19-JUL-2000) Arabidopsis thaliana Genome Center,
REFERENCE Department of Biology, University of Pennsylvania, 38th and
AUTHORS Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT On Jul 15, 2000 this sequence version replaced gi:2618677.
FEATURES
source location/qualifiers
1..101176
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="F21B7"
join(736..1038,1276..1713,1874..2119,2205..3032,
3117..3274,3483..3610,3725..4134,4313..4435)
/notes="unknown protein; similar to ESTs dbj|AV529515.1,
dbj|AV529266.1, dbj|AV54950.1, dbj|AV523761.1,
dbj|AV523761.1, dbj|A1996703.1, and dbj|AV546743.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.1"
/protein_id="AAF86508.1"
/db_xref="GI:9280639"
/translation="MOPLPARGSDHEGFWMNHPDLLLVAGDETNGTGLGHFSQNGSL
ARGDSKAGDAINYPPTTVRFYSLRSHSVVYVLRPSVCMIRCSRRVAVGLANOI
CYVDATLENKFSVLYTPQVPRQCTRVNCGYGMVAPRILAYAKSSMTKTR
LSQTFSTSSLSFSSSSGSSFWARYAMESSQLANGILNGLMGYKLSKCYQDML
PDGSTSPASNALKVGGSDAENAGMVAKDLVSGALVSQFKAHTSPISALCFDP
SGTLTLLVAVCGNNINVFQIMPSRSHNAPGDLSEWESHVHLFKLHRGITSIVODI
CFSQSQOWAIISSKGTCHIFVLNSSGDAAFQCEGEPTRLPASLPMWFTQSLSS
NOQLSPSTVALSVSRIRYKVSFGMLNTVSNATTAATGKVFVPSGAAVAFHKSVTH
DOLNSPTNALEHLVATPSGHVVOHELLPSVCTESPENGLRVQKTSHVQVQEDDLR
KVPTIQWDCVCRSDMLETERLPKSITEKQYDLETVSNHLTSHEDACLSDMNSHPS
EDKYLKSCSPERSHCYLSNFVEVKVTGMLPWNQSKLSFHWMDSPRDSSTGGEF
EIKVPAHELEIKQKLLPVDFHFTKATLEDRFSMKYCHTSATGSHQVNGYKTEKNN
INCHSGPGSTESAESSBEGSTKQENLHSDHMSNISKSLPLPYTVNGIYKTEKNN
ANGMEKPTVAKLTETRTINGTTPPTLTDVSNQEMLSCTGKPMGFGFALHEHC
KAVADPKEEHLKKKLDVNTVHNLVNNNTEKLGQDKMWNQSVLNAFGVCLYKDFL
RVGNPNFSFPLLLYKILRK"
complement(4838..6223)
/notes="hypothetical protein"
/codon_start=1
/evidence=not_experimental
/product="F21B7.2"
/protein_id="AAF86541.1"
/db_xref="GI:9280672"
/translation="MASCIOQLFTHLHIPVTINQOFLVHPSPPTPANOSPHLSLYLS
NLDDIICARVFTPSVYFPTNNRESFVLKRODALSEVLVYVPLSGRLREVENGLK
EVFGREQGVLMVANSNDMLADGLDTPNPANWPLIFRNPGEAYKILEMPLLIAQ
PVMVKYPHEHLYPIERSNLNLSMDTKPLOKCYRISKFEQCRVKSIAQEDPTLCKN
TFDAMAHAHWRWVKALDKVPLNLRLETSVNVTRLETKLRKGFYGNVCLACAM
SSDINSLSTKTRLVQADRLVRSVDYLRSMVDYVDVKKRKLREFGGLTKTQWTR
FENYETADGCMGKPVVAGPIDLRTPQCVLLPQGGVESGNDQSMVVCCLCPPTAVHT
FTRLNSLDHK"
complement(join(6700..7357,7482..7884,8620..8839,
8920..9241,9758..10380))
/notes="unknown protein; similar to ESTs gb|T04168.1,
gb|T21964.1, and gb|AW004462.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.3"
/protein_id="AAF86540.1"
/db_xref="GI:9280671"
/translation="MGHDSFYLVLRCLRCGLIAMLQICALQKKERRSKMESSDRS
SOAKAFDETGTGKLVASGIEKIPAMFHTPDTLTSLKOTAPPQOOLTPTVDLKGQ
SMDLISRRSVVEKIGDAERMGFQVNVHGISVEVMERKEGIRRHQDPEVKKRFY
SRDTRDLYYSNIDLHTCNKAANWRDTLACYMAPDPKQDLPAVCGEIMMEYSKOL
MTIGLEFLLSEALGLNPMLKDMGCAKSHIMFGQYPPCPQDPLTGLISKHTDFS
ITILLQDNGGLQVTHDQVSPVPGALVINIGDQLLISNDKFTSAERHVIANGS
SEPRISMPCEVSTFMKPNRIYGIKELSEQNPARYDLTITFEFNSFQSLRITRE
IPALFRATPATLASLSPPPKHLITPTVDLKASVVEKIGEAEEKWGLRHLNVHGYT
VEVERMIOGIRGFHEQEPAKKRFYSRDTRDLYVFNHDNLONSEAKSRDRLGCVT
APPRPELDLPVCGEIMLEYKSEIMSLGERLPELLSEALGLSHLKLKMDCKASQYM
VGQHPYCPQDPLTIGINKHTDISFLTVLQDNNVGGVQFHEQYHWDVTVPALYIN
IGDFLVQVKIMLLLRSCFISCMINGLILFYQLQINDKFIKSAERHVIANGSEPT
SVAIVSTFEMRAYSRVYGPVKILLSAENPAKYRDCLTETESTFSSKTLDAKLHFEK
I"
10833..11348
/notes="unknown protein; similar to EST emb|Z30724.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.4"
/protein_id="AAF86509.1"
/db_xref="GI:9280640"
/translation="MRGFSSVSSKWLTSVEIGVFLGSTTWLDAVAEADAVDMVFS
ACLMVKNKDLRLWLLRVPVAGLSEPPFSLRFLSPSIRIEGCFAGKFSDLRFRWSE
AWKCVVYLRGWLDDQIGWRVGNWVFSRVFVAGPSSVLRLDRODSTANDLVPVVV
LQSSGFCIRRY"
join(11965..12033,12173..12303,12448..12519,12596..12685,
12756..12915,12985..13041)
/notes="similar to ATH1 dbj|BAA37110.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.5"
/protein_id="AAF86510.1"
/db_xref="GI:9280641"
/translation="MNTIVAOQLQRFODYIVSLYQOGLDNOFSELRLKQDECTPDF
VAVSLYFPDCSKLINTMSISLERDNDVKQDVGSHOLKGSSSSLCFVLSVGARR
VNVVCSIFKCCCDVQNRGYSNHNCDNDINWHRCLRCLQQVDYIEYKMLTKLQDLF
NVKSGCENLVLVLLLEKQILQAGGTIPQVDIN"
16495..17682
/notes="similar to protein kinase pir|A57676; similar to
ESTs gb|A1993651.1, dbj|AV538995.1, dbj|AV522534.1, and
dbj|AV538995.1"
/codon_start=1
/evidence=not_experimental
/product="F21B7.6"
/protein_id="AAF86511.1"
/db_xref="GI:9280642"
/translation="MRFTKLWCLMELLRFGEFTEALLDPVDFLAQLAIRKSLDDLPG
SKFESWDFTSDCGAGVYCNKGDKVISLNLGDRPAGSPGLSRIDRPAIKLSALTLE
SIVPGRIMGALPATISOLKDLRLAISRNFIISGEIPASLGEVRLTDLDSYNOLTGT

```


/db_xref="GI:10187185"
/translation="MAAIENRVLTKHEVTKVTFVNPNSKTSFTLIDLTYDFWFVK
NPVPERLFFVEMDLTWDLFNSLTLKHLSTLLHYPLAGHIMWLDAAKPAVY
YFPDNGVSFAVAEWSSECHAGFHLSGNGIRQAVEFHLVPOLSTDDKAEVIAQ
ITLFPNPGESIGVSSHAILDGLKSTSLFKSWAYLCKQLCHLPCLSPELTLLDRT
VIDPQTQDMLQKWWVSDNSDPQKIRSLKVLFPJLDSLNLKVRATFELTREDIT
KLRRVNHQJLSKSKSKQVRLSTFVLTAYFCMAKAKLAKAKAEAAENDEIKN
IIIVGTADYRSRDLPPPLNYFGNGRHCEATKASDFVQENGVAFAEMLSDMVGI
DADAIEANDKVSLEILKLEKAMIFSVAGSTQFDVYSGDFGWRPKKVEIVSDRTQ
AISLABRRDGGGVGEVGVLEKQOMEVFSVFADGLKNDLV"
BASE COUNT 464 a 344 c 354 g 486 t
ORIGIN

Query Match 4.8%; Score 81; DB 9; Length 1648;
Best Local Similarity 46.6%; Pred. No. 4.9e-10;
Matches 643; Conservative 0; Mismatches 660; Indels 76; Gaps 9;

Qy 80 GTTATFCGTACCGGTAAACATTCTTCGATATCCCTCGTGTGCACTTGAATAAGATGCAATC 139
Db 141 GTTATCTCGATCTCACTATTTCGACTTTTCTGCTTGAAGAATCTCTCTGTGGAACG 200
Qy 140 CCTCTGTTTACGACTTTCGTACCCCAAGAACACATTTCTTGACACTG---TTATCCC 196
Db 201 CCTCTCTTCTATGAGATGACTGACTTGACGTGGGATTTATTTCAACTCGGAGATCCTCCC 260
Qy 197 TAATCTTAAGGCTCTTTGCTCTCACCTCTAAACACATCTCCGCTTACGGGAATTT 256
Db 261 AAAGCTGAAGCACTCCCTTCTTCACTCTCTTCACTTACCTCCCTCTGCTGTCACAT 320
Qy 257 GTTGATGCCGATCAAAATCGGCGAAATCGGAAGTTTCAGTACTCCGCTGATGAGGGCA 316
Db 321 CATGTGGCGCGTGA---TGCGCAAGACCTCGCCCTCTACTACTTTCCCGCAACAAAGCA 377
Qy 317 CTCGATACTTTGATCGTTGCGGAGCTGACAGGATTTTCACTACCTTAAAGTCACTCA 376
Db 378 CGGCTTTTCATTCGAGTTGCTGATG---GCTTCCGAGTGCCACGACGCGTTCCA 431
Qy 377 ACTGTAGATTCCAAATGATTGCGGCTTTTATGTTATGCGACGGTTATAAGGAC 436
Db 432 TCACCTCTCGGCAACGGAATCCGCAAGCAGTTGAATTTTCATCTCTGTGCCCCAGTT 491
Qy 437 CATGCAAGACTATAAAGTATCCCGCTGCTAGCGGTGCAAGTACCGTTTTCCTAACCG 496
Db 492 GTCGCTTACGACGACGATAAAGCTGAGGTAATTCGCAATCAACATACACTGTTTCCGAATCA 551
Qy 497 TGGCATAGCGCTGCTGACGGCACATCTCAATTCGACATGCTAAAGCTTTTGTAAAT 556
Db 552 AGGCTTCTCAATGTTGTTTCATCTCACCATGCAATCTTGTGGAANAATCTCGACCTT 611
Qy 557 GTTCATCAATGCTTGGGCTATATTAACAAATTTGGGAAG-----ACGCGGACTTGT 610
Db 612 GTTCTGAAATCTTGGGCTTATTTGTGCAACAATTAACAATATGCGCATCCCTTGT 671
Qy 611 GTCGCGGAATCTTTCATCTTTCGATAGATGCAATCAAAAGATC----- 657
Db 672 GTCACCTGAACCTCCCTCTTCGACCGGACTGCTCAAAAGATCGACAGGTCAGGA 731
Qy 658 -----TGATGGCTTAGAGGAACATTTTGAAGCAATGCAAGAT 698
Db 732 CATGCTGCAATGAATAAGTGGTGTGCGGTGCGGATAATTTTCGATCCCAAGAAATACG 791
Qy 699 GTTCTTGAATGTTCTCTAGATTTTGGGAAGCAACCCCTCGCATCAACAAAGTACGAGCT 758
Db 792 GAGCTTGAAGGTTTACCATTTCTAGACTCTGAGTCTCTGAA---CAATGCTCGGAGCC 849
Qy 759 ACATATGCTCTCTCCCTTGTGTAATCCGAAGCTTAAGAAAGTACTGAATCTCAGA 818
Db 850 ACATTTGAGTTGACGCGTGAAGATATTACGAACCTCAGGCACAAAGGTTAATCATCAGTTA 909
Qy 819 GGATCCGNAAC-----CACAACTAGCTTAAGCAGGTTCAACATGACGCTGTGGATAC 869
Db 910 TCAAAATCATCAAAATCAAGCAAGTTGCTTTATCAACTTTTGTGCTTACATGACTTAT 969

Qy 870 GTATGGACATCGATCGTCAAAATCAAAAGATGAC-----GTCGTATCAGAGGAA 917
Db 970 GTGTTTGTTCGATCGGTAAAGCTAAATTTACCAAAAGCCAAACTGAAGCTGAAGCTGCA 1029
Qy 918 TCATCGAAGCAGCAAAATGAGCTCGAGTACTTACGTTTACAGCGGATGCGGAGGACTT 977
Db 1030 GCAGTATGATGAATAATTAATAATTTATTTGCGGATTTCACTGGGATTTATAGGACCGT 1089
Qy 978 CTGACGCCCGCTGCGCCCTAACTACTTTGSCAACTCTCTTGGTCTATCGCTTGCAAAA 1037
Db 1090 TTGATCTCTCAATTCACCTTAATTTTGTGTAAGTCAATGGAGACATTTGAGACT 1149
Qy 1038 GCACACATAAAGAGTTAGTTGGGATAAAGGCTTCTTGTTCAGTTGCACTTGCAGTATGGA 1097
Db 1150 GCAAAAGCAAGTATTCGTTTCAAGAAATGCGGTGCTTTT---GTTGCAAGATGTTA 1206
Qy 1098 GAAGCCATTGAAAAGAGTTTGCACAAAGAAAAGCGCTTCTTGCAGATGCAAAACTTGG 1157
Db 1207 AGTATATGTTCAAGAGGATCGATCGGATGCCATTGAAGCCAATGATGATAAGGTTTCA 1266
Qy 1158 TTATCGGAATCTAATGGAATCCCTTCAAAAGATTTCTCGGATACCGGATCGCCTTAAG 1217
Db 1267 GAATATTTGAAATTTCTGAAAGAGGAGCAATGATTTTCTGTGCTGCGTCCGACCCAA 1326
Qy 1218 TTCGATTCGTATGTTAGATTTTGGATGGGAAAGCCTGCAAAATTTGACATTTACCTCT 1277
Db 1327 TTTGATTTTACGGTTCGATTTTCGGTGGGAGGCCCAAGAGGTGAGATTTGTGCA 1386
Qy 1278 GTTGATATGCAAGATTTGATTTATGATTCAGTCCAGGATTTTGAAGAGGTGAGG 1337
Db 1387 ATAGTAGACACAAGCCTCTCTTTGGCAGAGAGAGATGGAGGAGCGGCTTGAG 1446
Qy 1338 ATTGAGTATCATTTGCTTAAAGTTCATATGATGATTCATTTGCAAAATCTTTGAAGAGG 1396
Db 1447 GTTGAGTGTGTTTAGAGAGCAACAATGAGGTTTTTGAATCTGTATTTGCTGATGG 1505

RESULT 11.

AP000606/c 82360 bp DNA PLN 27-DEC-2000
LOCUS Arabidopsis thaliana genomic DNA, chromosome 3, pl clone: MTO24.
DEFINITION AP000606.1 GI:6045161
ACCESSION AP000606.1
VERSION 1
KEYWORDS Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui pl clone:MTO24.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (sites)
AUTHORS Kaneko, T., Katoh, T., Sato, S., Nakamura, A., Asamizu, E. and Tabata, S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 3. II. Sequence features of the 4,251,695 bp regions covered by 90 Pl. TAC and BAC clones
JOURNAL DNA Res. 7 (3), 217-221 (2000)
MEDLINE 20363099
REFERENCE 2 (bases 1 to 82360)
AUTHORS Kaneko, T., Katoh, T., Sato, S., Nakamura, Y., Asamizu, E. and Tabata, S.
TITLE Direct Submission
JOURNAL Submitted (13-OCT-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934)
COMMENT Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see <http://www.kazusa.or.jp/kaos/cgi-bin/agd-graph.cgi?c=MTO24>
Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory).

gene_id:WTO24.6"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01192.1"
/db_xref="GI:11994163"
/translation="MRRNSVSIYLQELSRVEVNVKSSIGKNKYADLGSNEDEMRDYATI
VMGELSGQTEQNEVILEYVDVDDTQDPEIISRRDORDDDRGDEYVEPPPAVSSQFKKEW
EDSI¹SLTLEHVEPSRAALHEVWDGKAFANSEYGYIKKSDKERRYNNKPPKTRISKNT
RMRKRGITVEALVHDDT²FCGLLEHTPAPVKIMKLVOTKGLGV³LSYSGTALRGROAVC
DLKGADESVKIDNICYMLKMLKNVSEVTYLIKLEDSQVGFLE⁴TALGASLEGFOAMRK
VILDTATHLKKYSYMSYRSRFFCICIBERYNNKSRITYLYLPDRSQNDIDDDIKALKLVPL
PRKKGRGYKVLRFPTSGRRRPPKROTQNKRAKQSLQWLLFGFHSI⁵
complement(join(24188,24189,24188,24334,24547,24687) .24774,
24898) .24962)
/note="emb|CAB61970.1
gene_id:WTO24.7
strong similarity to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01193.1"
/db_xref="GI:11994164"
/translation="MVVEPTKLSRKNEKDAQEPASTDPAGMSGQLTLELVDYGCSSRV
KRLCGKHFFAPYIRDEAAGLNDHLIT¹SRHHKTFELKVSQHL²SKLVGKGVRLGEYNST
ELMRLLPGPTVYMSVEDSRSDGMDREKSRKKDKC³IRTHPIAHGNCNITGKGR
KHQEWY⁴NIHHGTSYSLRSPSCY⁵ILLSHGETSRVRSRPNQDQNRSPRWPMDKTR
DVGHASKPRADGRRTSASLR⁶
join(25891,26085,26155,26349,26412) .26534)
/note="contains similarity to Mutator-like transposase
gene_id:WTO24.8"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01194.1"
/db_xref="GI:11994165"
/translation="MDRLCESDPWDEMKSAKRM¹QOLEEVAWMEIPI²ILPCGCGRTL
DIISEKDGKGRYECTDYKNDGLHIQKLDWKNKVEBNRLRQVNDHHQK³QSLSEY
SNFPLYSLYR⁴
complement(join(26878) .27120,27210) .27491,27693) .27854,
27950) .28175,28419) .28702,29026) .29487))
/note="emb|CAB77996.1
gene_id:WTO24.9
strong similarity to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB01195.1"
/db_xref="GI:11994166"
/translation="MAPRTIPLPKRKGVAGGTPKIDVVT¹KATERTPTTTEPSATE
QNPVILEGGVEEPIPT²IVVSEGGSDNNBENS³DKDSESESESESESESESESESESE
EKKEEKEEGNVTGGSSENDSTLGEYS⁴SDKMNMDETA⁵VENOV⁶EIPASMEIDET
KALILSPMPFPEYVKK⁷ILKISRCY⁸IEHVLT⁹FDKLEPENKSEREFQNHPSFQHI¹⁰
FHMPPNRILMKMLPRTAIRERKEAC¹¹I¹²IAPIKTGPKAPQVEBFCGLKAVSDLT
FCRNFGWGRFYDYL¹³MGIT¹⁴SHVNFHNSGVNTN¹⁵NEKYVMSVGLFCGLMELAFEAIPOL
RTFMEDIAAGBGCRCMCKYR¹⁶FKKNHLKGP¹⁷LDIT¹⁸YAEGLT¹⁹ORRKESLR²⁰IRPDEYI
QDQI²¹QALQITPKAGAD²²IPAVGGQELPS²³SVYVMEI²⁴NLKKMVDQ²⁵NEILITMIS
DLDRVSELEAFKDEKQAEERKNQEMQEKDGP²⁶IEKDGQV²⁷QRKDGQEDGVEPKKED
QERDGVPEKFEKQKQKGGPSK²⁸PEANFT²⁹TRITRTKN"
join(33399) .33608,33762) .33830,34025) .34098)
/note="gene_id:WTO24.10
unknown protein"
/codon_start=1
/pseudocodon
/evidence=not_experimental
join(34168) .34242,34530) .35934,36128) .39479)
/note="gene_id:WTO24.11"
/codon_start=1
/pseudocodon
/evidence=not_experimental
/products="retroelement pol polyprotein-like"
39663) .40357
/note="gene_id:WTO24.12
unknown protein"
/codon_start=1
/pseudocodon

```
/evidence-not_experimental
Join(49318..49683,49880..49959,50068..50400,50660..50698..50779..50980)
/feature="contains similarity to Mutator-like transposase
gene_id:MT024.13"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB01196.1"
/db_xref="gi:11994167"
/translacion="MKRNDSTKLVSPFSGRKLPRNVIRNDELCYFEDV
NEEFPSILHVESNWRNENRPFKILREDLVRONVNNEDSLIVGTDEIGAVLAM
YVGPSEVVNENKDIYLEENICGGMSPYSIRKDSFMSIVKYNIRAKRFQVQVSTAW
KGRKAANDIRGPEEFVDFYDYYMLQKMHESGYSVEDEENKYNLFALGSSIT
EGFRLMRKVIILDSHTKMTAYGGLIVATTQDPVHRCASPECARAYTAEFURCYF
GGXYNRDTSVSVNGVLEKVRNYSLLQLIDAIVGKIAEFWFAKHKRSLLELIPSGY
LDPPD"
CDS
Join(58702..58758,59228..59328,59533..59839)
/feature="gene_id:MT024.14
unknown protein"
/codon_start=1
/evidence-not_experimental

Query Match 4.2%; Score 71.8; DB 13; Length 82360;
Best Local Similarity 46.2%; Pred. No. 1.7e-07;
Matches 604; Conservative 0; Mismatches 627; Indels 76; Gaps 8;

QY 187 CTGTTATCCCTAATCTTAAGGCTCTTGTCTCTCACTCTAAACACTACGTTCCGCTTA 246
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16389 CTCTCATCTTAACCTTAAGTCTCTCTCTCAGCCACGCTCTCCACTATCTCCACTCG 16330

QY 247 GCGGAAATTTGTTGATCGCGATCAATCGGGGAAATGCCGAAGTTTCAGTACTCCCGTG 306
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16329 CGGGCGGCTCGTGGGATTCAATCAAAACCAA---CCATCCATTGCTATTCTCCTG 16273

QY 307 ATGAGGCGACTCGATACTTTGATCGTTGGGAGTCTGACCAGGATTTTGACTACCTTA 366
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16272 ACGAAGAAGTCCGCTGATCTCAGCGAGTGTGATCTTAACGGTGACCTTTCCCATCTTT 16213

QY 367 AAGGTTCATCACTGTAGATTCCAAATGATTTGCATGGCTTTTATGTTATGTCACACGG 426
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16212 CGGAGATGAACCTCGTCCGCCACGAGTTTCATCTTTAGTCCCTGAGTTACC----- 16156

QY 427 TTATAAGGACCATGCAAGCATATAAAGTATCCCGCTCGTAGCGTGCAGTAACCGTTT 486
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16157 -----GGTTCCGAGGAATCAGTTCGGTCTCGCGTTCAGGTAAAGTTT 16111

QY 487 TTCCTAACCGTGGCATACCGTGGCTCTGACGGCACATCATTCATTCAGATGCTAAAA 546
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16110 TCCCAACCAAGGATTTTCTCTCGCGTAAACGGCACATCATTCGCGTTTATGATGGTAAAA 16051

QY 547 GTTTGTATGTCATCAATGCTTGGGCTATATTAAACAAATTTGGGAAAGACGGGACT 606
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 16050 CAACAGCATGTTTCTCAAGCATGGG-----TCACACTGTAACAAAGAACGAAG 15997

QY 607 TGTTCGCGAATCTTCTCCATCTTCGATAGATCGATTAATCAAGATCTGTATGCC 666
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15996 CTCCTGCCACAGATCTAGTACCGTCTTTAGATCGTATCATTTTCAAGACCAACGGGAC 15937

QY 667 TAGAGAACATTTTGGAAACGAATGCAAGATGTTCTTGAATGTTCTAGATTGGAA 726
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15936 TCGAACAAACTCTTGAACCGGTGGATATCAGCTTCCAAACAAACCAAGCTTAAAC 15877

QY 727 -----GCAACCCCTCGATTCAACAGGTACGAGTACATATGCTCTCCCTTG 777
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15876 TGTTCCTTCCAAATAATTTGATCCGATATCTCCGAGTACGTCAGGTAACTCGAG 15817

QY 778 CTGAATCCAGAACTTAAGAACAAAGTACTGAATCTCAGAGGATCCGAACCGCAATAC 837
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15816 AAGACATCAAGAAGCTA---CGTGAACGAGTGGAGACCGAGTACACAGCTAAGCAGCTCC 15760

QY 838 GTGTAACGAGTTCACATGAGTGTGGATACGATGATGATGATGATGATGATGATGATGAT 897
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15759 GTTTCGAGCTTTTGTCTTACTTACGCTTACGTTATTAACATGATGATGATGATGATGATG 15700
```

```
QY 898 ATGACGTGTTATCAGAGGAATCATCGAACGACGACGAAATGAGCTCGAGTACTTCACTTTTA 957
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15699 GAG-----CTGACCCGACCGGTTTCTGTGCGTGTGTTG 15664

QY 958 CAGCGGATTCGCGAGGACTCTGACGCCCCCGTGTCCCTAACTACTTTGGCAACTGTC 1017
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15663 CATCAGACTTTCCGTCACGTTTAAACCTCCGTTACCCCGACGTTCTTCGGGAAGTGA 15604

QY 1018 TTG-----CGTCATCGGTTGCAAAACACACATAAAGAGTTAGTTGGGATTA 1066
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15603 TTGTCGGTTCGGAGATTTCGACGCTGAAAGCAGAACCAATATTTGGAAGAGAGAGAA 15544

QY 1067 AGGGCTTCTGTTGCGAGTTGAGCTTATTTGGAGAACCCATTTGAAAAGAGGTTGCACAA 1126
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15543 AAGTTTCATCAGCGCTGGAGACTTTGACCGGATGGGTCAACGGATGTGTCGGGAGA 15484

QY 1127 AAAAGCGTCTTCGAGATGCAAAACCTTGGTATTCGGAATCTAATGGAATCCCTTCAAA 1186
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15483 ATATAGAGAAGAAATATGCTTTTACCGTTTGAAGCGTTTAAAGAGGATGGAACCGGAAGC 15424

QY 1187 AAGATTTCTCGGATACCGGATCCCTAAGTTCGATTCGATTCGATTCGATTCGATTCGAT 1246
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15423 AATGATATCGG--TTGCTGGTTCGACCGGTTAGGGATTTACGGGTCTGATTTTGGGTG 15366

QY 1247 GGGAAAGCTGCAAAATTTGACATTAACCTCTGTTGATTCGAGAAATGATTTATGTGAT 1306
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15365 GGGAAACCGGTCAAGGTAGAGATTGTCAGGATTCATGAAGGATGCTTCGGTTTCATGTC 15306

QY 1307 TCAGTCCAGGATTTTGAAGAGTGTGAGATTCGAGATTCGATTCGATTCGATTCGATTCAT 1366
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15305 GGAGAGTGGAGATGGAAGTGTGTTGAGTGTGTTGAGTGTGTTGTTGTTGTTGTTGTTGTTG 15246

QY 1367 GSATCATTTCGAAAATCTTTGAGAGAGGCTTTTGCTTTGTCATAGTCTCTTTAATA 1426
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15245 GGAAGCTTTGTTGTTTATTCCTCCATCGGTTGGAATAAAATGAGAACGTTTGGTTTC 15186

QY 1427 GAACCATATTTCTCAATAAAGTACCAAGTCTTTAGTACACACTAC 1473
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 15185 TAGTTTATGTTGTTTATTTTATCCACATCATCTTTAATTCATTAC 15139
```

```
RESULT 12
AC074226
LOCUS AC074226 341064 bp DNA HTG 15-SEP-2000
DEFINITION Arabidopsis thaliana chromosome 1 clone IGF-F27B9, *** SEQUENCING
IN PROGRESS ***, 5 unordered pieces.
AC074226
VERSION AC074226.3 GI:10140624
KEYWORDS HTG; HTGS_PHASE1.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 341064)
Lin, X., Kaul, S., Town, C.D., Benito, M., Creasy, T.H., Ronning, C.M.,
Koo, H., Fujii, C.Y., Utterback, T., Barnstead, M.E., Bowman, C.L.,
White, O., Nierman, W.C. and Fraser, C.M.
Arabidopsis thaliana 'IGF' BAC 'F27B9' genomic sequence near marker
,
Unpublished
2 (bases 1 to 341064)
Town, C.D. and Kaul, S.
Direct Submission
Submitted (20-JUL-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
On Sep 15, 2000 this sequence version replaced gi:10086523.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 5 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
```


* This record will be updated with the finished sequence
 * as soon as it is available and the accession number will
 * be preserved

1 14022: contig of 14022 bp in length
 * 14073: gap of unknown length
 * 14073 35113: contig of 21041 bp in length
 * 35114 35163: gap of unknown length
 * 35164 86858: contig of 51695 bp in length
 * 86859 86908: gap of unknown length
 * 86909 170536: contig of 83628 bp in length
 * 170537 170586: gap of unknown length
 * 170587 341064: contig of 170478 bp in length.

FEATURES
 Location/Qualifiers

1..341064
 /organism="Arabidopsis thaliana"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="IGF-F27B9"

BASE COUNT 110402 a 59862 c 59862 g 110479 t 459 others
 ORIGIN

Query Match 4.2%; Score 71.8; DB 75; Length 341064;
 Best Local Similarity 46.2%; Pred. NO. 1.9e-07;
 Matches 604; Conservative 0; Mismatches 627; Indels 76; Gaps 8;
 QY 187 CTGTTATCCCTAATCTTAAGGCTCTTTGCTCTCACTTAAACACTACGTTCCGCTTA 246
 DB 65157 CTCTCATCTCTAACTTAAGTCTCTCTCAGCCACGCTCTCCCACTATCTCCCACTCG 65216
 QY 247 GCGGAAATTTGTTGATGCGGATCAATCGGCGAAATGCGGAAGTTTCACTAGTACTCCCGTG 306
 DB 65217 CCGCGCCCTCGTGTGGGATCAATCAAAACCAAA---CCATCCATGTCATTCTCTCG 65273
 QY 307 ATGAGGCGCACTGATGAATCTTTGATGTTGCGGAGTCTGACCAAGGATTTGACTACCTTA 366
 DB 65274 AGCACAAGATGCCGTGTATCTCAGCGTAGCTGAGTCAACGCTGACCTTTCCCATCTTT 65333
 QY 367 AAGGTATCACTGGTAGATTTCAATGATTTGCATGGCCCTTTTATGTTATGCCACGGG 426
 DB 65334 CCGGAGATGAACCTCGTCCCGGACCGGAGTTTCATCTTTAGTCCCTGAGTTACC----- 65388

QY 427 TTATAGGACCATGCAAGACTATAAGATGATCCCGCTCGTAGCCGTGCAAGTAAACCGTTT 486
 DB 65389 -----GGTTCCCGAGATACGTCGCGTTCTCGCGTTCAGGTAACGTTT 65435
 QY 487 TTCCTAACCGTGGCATAGCGTGTCTGACGGCAGATCATTAATGTCAGATGCTAAAA 546
 DB 65436 TCCCAACCAAGGATTTCTCTCGCGGTAAACGGCAGATCATGCGGTTTATGATGTTAAAA 65495
 QY 547 GTTTTCTAATGTTTCATCAATGCTTGGGCTATATTAACAAATTTGGGAAGACGCGGACT 606
 DB 65496 CAACGCCATGTTCTCAAGCATGGC-----TCACAACTGTAAACAAAGCAAGAAG 65549
 QY 607 TGTGTCGCGGAATCTTCTCCATCTTCGATAGATGATCAATCAAGATGCTATGTCG 666
 DB 65550 CTCTGCCACAGTCTAGTACCGTCTTTAGATCGTATCATTTTCAAGACCAACCGGAC 65609
 QY 667 TAGAGAAACATTTGGACAGAAATGCAAGATGTTCTTGAATGTTCTCTAGATTTGAA 726
 DB 65610 TCGACAAACATCTTGAACCGGTGATATCAGCTTCCCAACAAACCAAGCTTAAAA 65669
 QY 727 -----GCAACCCCTCGATTCAACAAAGTACGACCTACATGTCCTCTCCCTTG 777
 DB 65670 TGTTCCTTCCAAATTAATTTGGATCGGATATTTCTCGGATCGACGTACCGGTTAACTCGAG 65729
 QY 778 CTGAATCCAGAGCTTAAAGAAAGTACTGTAATCTCAGAGGATCCGAACCAACATAAC 837
 DB 65730 AAGACATCAAGAAGCTA---CGTACAGAGTGGAGACCGGATGACAGCTCAAGCAGCTCC 65786
 QY 838 GTGTACAGAGCTTCAATGAGCTGTGGATAGCTATGAGACATGATGTCATGTCATCAAAAG 897
 DB 65787 GTTTGTCGACTTTTGTTCATTACTTACGCTTAGCTTATAACATGCATGCTGGAAGATCGCTG 65846

QY 898 ATGACGCTCGTATCAGAGGAATCATCAACGACGARAATGAGCTCGAGTACTTTCAGTTTTA 957
 DB 65847 GAG-----GTGACCCGACCCGGTTTGTGTCGGTTTTG 65882
 QY 958 CAGCGGATTCGCGAGGACVTTCTGAGCCCGCCCTGTCGCGCTAACTACTTTGGCAACTGTC 1017
 DB 65883 CATCAGACTTTCGGTCACTGTTAAACCCCTCCGTTACCGCGGACGTTCTTTCGGGAACGTCA 65942
 QY 1018 TTG-----CGTCATGCGTTGCAAAAGCAACACATAAAGAGTTAGTTGGGGATTA 1066
 DB 65943 TTGTCGGGTCGGGAGATTTCGACGTGAAAGCAGAACCAATATTGGGAAGAGAGAGAA 66002
 QY 1067 AGGGCTTCTTGTTCAGTTGCGAGCTATTGGAGAAGCCATTGAAAGAGAGCTTGCACAACCA 1126
 DB 66003 AAGGTTTTCATCAGCGCGGTGGAGACTTTGACCGGATGGGTCAACGGATTTGTGTCGGAGA 66062
 QY 1127 AAAGGCGTTCTTCAGATGCAAAACCTTTGTTATCGGAATCTAATGGAATCCCTTTCAAA 1186
 DB 66063 ATATAGAGAAGAAATATGCTTTTACCCTTTTAAAGCGTTTAAAGAGATGGAACCGGAAGGC 66122
 QY 1187 AAGATTTCTCGGATTTACCGGATCGGCTAAGTTTCGATTCGATATGTTGTTAGATTTTGGGATG 1246
 DB 66123 AATGATATCGG--TTGCTGGGTGCGACCGGTTAGGGATTTACGGGTCTGATTTTGGGTG 66180
 QY 1247 GGAAGCGCTGCAAAATTTGACATTAACCTCTGTTGATTTATGCAAGATTTGATTTATGTTGAT 1306
 DB 66181 GGAAGCGGTTTGAAGAGGTTGAGAGATTTGACGATTTGATAGGATGCTTCGGTTTTCATTGTC 66240
 QY 1307 TCAGTCCAGGGATTTTGAAGAGGTTGGAGATTTGAGATATCATTTGCCCTTAAGATTCATAT 1366
 DB 66241 GGAGAGTGAGATGGAAGTGGTGGTTCGAGTTGGTGTGTTGTTTGAAGAAAGATGATGT 66300
 QY 1367 GGATGATTTGCAAAATCTTTGAAGAGGCTTTTGCCTCTTTGTCATAGTCTCTTTAATA 1426
 DB 66301 GGAACGTTTGTGTTTCATTAATCTCCATCGGTTGGAATAAAAATGAGAAGCTTTTGGTTC 66360
 QY 1427 GAACCATTTTCTGCAATTAAGTACCAAGTCTCTTTAGTAACACTAC 1473
 DB 66361 TAGTTTATGTTGTTTTCATTCACATCATCTTTAATTCATTAAC 66407

RESULT 13

AB016892
 LOCUS AB016892 66237 bp DNA PLN 27-DEC-2000
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MXF12.
 ACCESSION AB016892 BA000015
 VERSION AB016892.1 GI:3449333
 KEYWORDS
 SOURCE Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MXF12.
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE 1 (sites)
 ASAMIZU, E., SATO, S., KANEKO, T., NAKAMURA, Y., KOTANI, H., MIYAJIMA, N. and TABATA, S.
 Structural analysis of Arabidopsis thaliana chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by seven clones physically assigned P1 and TAC clones
 DNA Res. 5 (6), 373-391 (1998)
 JOURNAL 99156233
 MEDLINE 2 (bases 1 to 66237)
 REFERENCE Nakamura, Y.
 AUTHORS Direct Submision
 TITLE Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)
 JOURNAL Address for correspondence: kaos@kazusa.or.jp
 COMMENT For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=MXF12

Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://compbio.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://grem1.n1.zool.lastate.edu/cgi-bin/sp.cgi>). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K15E6 and the 3' clone is K3K3.

FEATURES source

Location/Qualifiers
1. .6237

/organism="Arabidopsis thaliana"

/strain="Columbia"

/db_xref="taxon:3702"

/chromosome="5"

/clone_lib="Mitsui P1"

2613. .5234

/note="gene_id:MXF12.2"

/codon_start=1

/evidence=not_experimental

/product="receptor-like protein kinase"

/protein_id="BAB10824.1"

/db_xref="GI:10177545"

/translation="MIRHALLIPSILVSTPIVGEATSTVEPTDVFLENGCDTSNNVD
VSGNRTAENOKILSNLVNASTAQASVQSVQIPYMTARIFRSEYTFPPVTPG
SNRLIFYPTRTGSGQNAVSEFFSVKNGFTLNNFSAFLTAKSKQTEFIKEFI
IPVQYLNFTTSLDSLAFVNGEIVSPNRYFSGGDDVITNNGSSVDFFHENT
AFETVRLNVTGDSGFRWRWDDIELSESGISIPVDIKINTEKTPSYVA
PDVYATSRMGNAHPQONLNLTLFTVDAGFSYLVRLHCFEKLSEVNGEGORVF
SIENQOTAFVDMFGSGSWIPMWLDYTVIAGSGRRHDLRLDHPVLSINPKY
YDALNGVELKMDPGLNAGNPDPVSPDLIPNRPRIKNSHILPILAVVG
SLVLAFFVGVVLIMKKKKSPKSNSSCPLPHGTDSTNTPAKSLPADLCRRFSI
FEKSNATNDKLLIIGVGGFSYKQIDGGATLVAVKREITSNOGAKEFEFELEM
LSKRLHVLSTLGYCDEENMLVYEPHGLTKDHLFRDKTSDPPLSMKRELEIC
IGARGQLYLTKAKTYI IHRDKITNNLIDENFVKVDFGLSRVGPSTASQTHVST
VWKTGYLDPEYRQVLTPEKSDVYSGVVLVLECCRPIMQSVPEQADLIRVVK
SNYRGTVQDQIIDSLSADITSTSLKFCFCAVRCVQDRGMERPMNVVWALEFALQ
LHFTAKKNDNVSLDLMPSGEGVTTTDDGDDLFSTRTGHVGKSTTTDDSVLVGDER
SGSSWGVFSINEPKAR"

6701. .7210

/note="gene_id:MXF12.3"

/unknown protein"

/codon_start=1

/evidence=not_experimental

/protein_id="BAB10825.1"

/db_xref="GI:10177546"

/translation="MINLEKEAKEQILIRDKRKIVHHVNTTLLVKIYEASAIIDRRH
LFLKPYGPNVYFIKAWLIDDTKEYGVYVVKDSFKVNQPKIPLNFPAGYLVKELLK
GMSYRDPGTSNRVTVMWGRAKIRLSPLMTSYKERPISSGIDLVLGLNSDQCQVKKGYLT
AKLHRSAD"

7670. .10111

/note="gene_id:MXF12.4"

/codon_start=1

/evidence=not_experimental

/product="receptor protein kinase-like protein"

/protein_id="BAB10826.1"

/db_xref="GI:10177547"

/translation="MNCNVLFLLSVLVSVTAGVTAAYHPTDVFLENGCDTSNNVDNSG
RNMTVFSRQILSNLVNASTAQASVQSVQIPYMTARIFRSEYTFPPVTPGSI
RLRYFYQVQKSDAVNPFVKNVNGFTLLRNFNADSTVQASIPLSNLSIKKEFIIPV
HOTLNLFTFTPSKLLAFVNGIEIVSMRDFYSGGDFNVLNVNSDQIDNSTAFE
SVRLHVGQIVNEVDSDSGMFRWLSDSDSGSISVNVVPGVINKINTEKTPAYVAPYD
VYATSRMGNSNLMNLGMFLTVDAIGNYLVRLHCFETLPQVTRAGQRFVIFVED
KMAKKTVDVIRLSGGPRIPRMVLDVSVYGVFESGMIOPELRLDLVPLKDTNQTYDAIL

SGVEILKLNDSGNLARPNPELVSTDDPDDSNVTPPIKPKPHVILVILVVGSVIG
LATFIVIMLLIRQMRRKKKNSVIMFKLLKQIYIAELKQIISFSTFSGKGGFG
TVYRNLNGRTVAVKVLKDKNGDDFNEVTSQTSQSHVNIVLSLFCYEGSKRAI
ISEFLHSGLDQFISRNKSLTPNVTLYGIALGARGLEYLHVCGCKTRIVHPDKPON
ILLDDNCPKADVAGLAKCEKRESILSLIDTRGTIGYTAPEVYSRWYSGIKSHKSDY
SYGMVLDMIGARKVETTCNGSTAYFPDWIYKDLKLENGDQTWIIGDEINEENKIVK
KMLVLCWIRPCPSDRPPMKNKVMIEGSLDALELPPKPSRHSIHELVELSSLSLDG
QEAKEQTQIDTII"

CDS

10819. .13239

/note="gene_id:MXF12.5"

/codon_start=1

/evidence=not_experimental

/product="receptor protein kinase-like protein"

/protein_id="BAB10827.1"

/db_xref="GI:10177548"

/translation="MICFILVFSLVSATAPYKPDVFLINCCETDVPDNGHGT
WTOEKNILPKNSDNASFSVSVTNDFTLLQNFSAFLTAKSKQTEFIKEFI
LUTYPTYSKGFDAVNSVSVTNDFTLLQNFSAFLTAKSKQTEFIKEFI
LTLNLTFRPSNNSLAFVNGIEIVSMRDFYSGGDDVITNNGSSVDFFHENT
VHRLNVGGMVDENVDSGMFRWLSDSDSGSISVNVVPGVINKINTEKTPAYVAPY
VYSTRMGNAQDTYLNLFNLTLFTVDAGFSYLVRLHCFEKLSEVNGEGORVF
NOMAREMDVIRLSGGPRIPYLDRIYVSGESGPRDLRLDHPVLSINPKY
NGVEILKLNNSGNIILQNELAPNEPLSNLTPNHVTOQIKGKSSHLVKKIIVGVP
GTGLATFVVLMLWMRQMKRKRKEERVNFKLLNNIYIAELKQIISFSTFSGKGG
FGTVYGNLSNGRKVAVKVLKDKLSAEDEFNEVTSQTSQSHVNIVLSLFCYEGSKR
AIVYFLENGSLDQFMSRNKSLTPNVTLYGIALGARGLEYLHVCGCKTRIVHFDIKP
ONILLDNLCPKVSDFGLAKCEKRESVLSMDTRGTIGYTAPEVYSRWYSGIKSHKSD
VYSGMLVIDMIGARKEIVETVDSASSTYFPDWIYKDLKLENGDQTWIIGDELTKEK
EIAKKMIVVGLWCIQPCPSDRPPMKNKVMIEGSLDALELPPKPSRHSIHELVELSS
LSDGGBDV"

CDS

join(16412. .16498,16859. .16924,17003. .17038,17085. .17192,
17337. .17461,17541. .17640,17803. .17974,18073. .18191,
18285. .18395,18488. .18632,18712. .18844,18934. .19009,
19232. .19297,19385. .19492,19591. .19698,19873. .20070,
20164. .20373)

/note="gene_id:MXF12.6"

/codon_start=1

/evidence=not_experimental

/product="ABC transporter-like protein"

/protein_id="BAB10828.1"

/db_xref="GI:10177549"

/translation="MGNKLLITGGSKTHGSSYRDPQLQNGEDPKANGSBNGLND
LEHGVVIAENAGFRVFALICQYSSIMSOAKPDAGLVIGTIALIGTSLNLLPKF
GMIIDIVSDKVTPEQTESLIARNAVVIILVVGISCTIALRAWLFNSRERV
ARGKDLFRKHLHQAETADVTGELLSSRLSDTOIKNAATNLSEARLVTTALI
GURKMTSSKMLTLALVAVKQFGRVLRSLSHTTQAAAVAAASIESGCA
VVRVSPAKESYMSQYKSKVDETLKLGKQAVLGLFEGGLNAETLSVTVVSGA
YLITGSMVGTALTSILITVSGSSVSSLSYTTAMKAAGSRVQFLDRVSSMS
SSGDCPVGPNPDGVELNDVMFAYPSRPSHMLKGLISRLTPGSKVALVPGSGGKTT
IANLIRFYDPLKGLILLNGVSLMEISHOYLHKQIISVQEPILFNCVSEENIAYGFD
GEASFTDIENAKMANAHFIEAFPKYNTVVGGRGLRSGGQKQRIATARALTNP
VLLDDEATSDALDAEVLVDQAMDLSMAGRTVLVIAHRLSTVKTADCVASIDGVAE
KGTHDELLSLGIYTNLVRQLOSSSVTTL"

join(25348. .26359,26401. .26756)

/note="gene_id:MXF12.7"

/codon_start=1

/evidence=not_experimental

/product="anthocyanin acyltransferase-like protein"

/protein_id="BAB10829.1"

/db_xref="GI:10177550"

/translation="MVNEEMESLKVIVARVTPSNSDSESILTPTFDDLLWYKLH
AVERVIFYKLTASRPFDFSVIVPNLKTSLSSLSHYLPLAGLVMVEPLDKPKIVYT
PNDVAFYVAESNADSRLTGKPFTEYPLVPELHVSDSDASAVSFQVTLFPNQ
FCISVNAHVAFLDGKTTNFKSMARTCKNODSFLPDQLIPVDRTVVKDPMOLDKI
LNAAHVRVFTGCKKEPNPKSLKLLSPEIGDFDPYRTNLNTRBIOQLRERLKES
LVNHRVSKKRELRTFVIIVYALTCLIKARGDPSRPVGYGFAVDCRSLMWPPVPS
SYFETPMSEGFALAAARMVSDVEALDENVALKIPILTEGFTTILSPGTQLVSVAGSTR
FGYVLDGFGWREKPVVWSIDQGEAISFAESRDGSGGVGLFSLKHKHMDVLVDLLH
KGLN"

complement(join(27466. .27879,28046. .29073))

/note="gene_id:MXF12.8"

/codon_start=1

/pseudo


```
/evidence=not_experimental
/product="transposase-like protein"
complement(join(30804..30874,30947..31289))
/note="contains similarity to ac-like transposase
gene_id:MXF12.9"
/codon_start=1
/pseudo
/evidence=not_experimental
32432..33823
/note="gene_id:MXF12.10"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin acyltransferase-like protein"
/protein_id="BAB10830.1"
/db_xref="GI:10177551"
/translation="MVSSNIIEVARVTPSNSDSAESLSPLTYFDLIYYKLRAVERV
IFRYINTWRPFDSVYVPLNKLTSLSCLSHYPLAGKLWEPLDHPKPIVYSQNDV
SFSVAIVTDFSSLSGNPFSTELYPLVPAQSSDSASIVSFQVTLFPNGFCIGV
SAHVAVLGDKTTFMFKSWAHCKHQDFSLPDQLIPTDYDRYTIKSPSEKNVLEWR
SFTKILAGKKEPANPKSLNPFSEIGRDPVVRYYTLQLTREDIOTLERLKRREVSSSS
STSSKELRLSTFVIVSYVLVCIIRAGGEPHVPVGFVSDCRSLMNPPTPNVFGN
CIAGSRMLTAKMFMGEGLLAATMVSDSLEWDESFANKIPDFVAYATLPPETOL
ILVSGNSRGVYELDFNGRDPKVMVVISLSPNGISMAESRDQNGVEIGFSLKKHEM
DFLIDLHLCEUFI"
34363..35709
CDS
Query Match 3.8%; Score 64.2; DB 12; Length 66237;
Best Local Similarity 49.4%; Pred. No. 1.4e-05;
Matches 222; Conservative 0; Mismatches 223; Indels 4; Gaps 2;
QY 948 TTCAGTTTACAGCGGATTCGCCAGGACTTCTGACGCCCGGTGTCGCGCTAACTACTTT 1007
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26299 TACGGATTGCTGCGATTGTCGAAGCCTTATGGTTCACCGGTTCCATCGAGTTATTT 26358
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1008 GGCACACTGCTCGCTGATCGCTTGCAAGACCAACATTAAGAGTAGTTGGGGATAAA 1067
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26359 GGTAAATTGTTGATCTGCTTG-TTTAAATGTCGTTAACGGCAGAGACTTTTATGACTGAA 26417
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1068 GGGCTTTGTTGTCAGTTGCGAGCTATTGGAGAAGCCATTGAAAGAGGTTGCACAACGAA 1127
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26418 GAAGGGTTTTAGCTGCTCGGAGAAATGGTTAGTGATCTGTTGAGGCATG---GATGAG 26474
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1128 AAGCGGTTCTTTCAGATGCAAAAATCTGGTTATTCGGAATCTAATGGAATCCCTTCAAAA 1187
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26475 AATGTTGCATTGAAGATTCCAGAGATTTTGGAAAGGTTTACGACTCTTCCACCAGGAACA 26534
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1188 AGATTTCCTGGGATTAACCGGATCCCTAAGTTCGATTCGATGGTGATGTTGGATGG 1247
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26535 CAGGTTTGTCTGTTCCGGGTGACCCGGTTTCGGAGTTTACGGGTTGGATTTGGGTGG 26594
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1248 GGAAGCCTGCAAAATTTGACATTACCTCTGTTGATTATGCAGAAATGATTTATGTGATT 1307
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26595 GGTAGACCGGAGAAAGTGGTGGTTGTGTCANATGACCAAGGTGAAGCGGATTTCTTTTGG 26654
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1308 CAGTCCAGGGATTTTGAAGAAGCTGGAGATTGGAGATTGAGTATCATTTGCTTAAGATTATG 1367
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26655 GACAGTAGAGATGGAGTGCTGCTGAGACTTGGCTTCCCTCAAGAAACATGAATG 26714
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1368 GATGCATTTGCAAAATCTTTGAAGAAGG 1396
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 26715 GATGTTCTTGTGATTGCTTTCACAGGG 26743
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
RESULT 14
AC069470/c
LOCUS
DEFINITION Arabidopsis thaliana chromosome I clone IGF-F7F7, *** SEQUENCING IN
PROGRESS ***, 4 unordered pieces.
ACCESSION AC069470
VERSION AC069470.10 GI:12320593
KEYWORDS HTG; HTGS_PHASE1.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
```

```
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 100163: contig of 100163 bp in length
* 100164 100213: gap of unknown length
* 100214 112623: contig of 12410 bp in length
* 112624 112673: gap of unknown length
* 112674 116336: contig of 3963 bp in length
* 116337 116886: gap of unknown length
* 116887 136047: contig of 19361 bp in length.
* Location/Qualifiers
1..136047
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="I"
/clone="IGF-F7F7"
BASE COUNT 46028 a 22872 c 22213 g 44783 t 151 others
ORIGIN
Query Match 3.7%; Score 63; DB 74; Length 136047;
Best Local Similarity 45.1%; Pred. No. 3.2e-05;
Matches 436; Conservative 0; Mismatches 520; Indels 11; Gaps 5;
QY 460 CGCTCCGTAGCGTGAAGTAAACCGTTTTCCTAACCGTCAGTACCGTGGCTCTGACGG 519
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28857 CGGTTCTTTTGGCAAGTCACTTTGTTCCCAACCAAGGGTTTGTATTGGTATCGCTT 28798
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 520 CACATCATTAATTCAGATGCTTAAAGTTTGTATGTTTCATCAATGCTTTGGGCTATTA 579
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28797 CTCATCATCTGTTATGGACGGGAACAGTGGTTAGGTTTATAAATCATGGCTCACA 28738
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 580 TTAAACAAATTTGGGAAGACGCGGACTTTGTTCCCGGAATCTTCTTCCATCTTTCGATA 639
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28737 TTTGTAACATGCA-----GCCATGGATTACCGGAGGATTTGACTCCGGTTTAGATC 28684
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 640 GATCGATAATCAAGATCTGTATGGCTAGAGGAACATTTTGAACGAATCAAGATG 699
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28683 GTACGGTTATCAATGTCCCTGCAAGTCTTGATGCAAAAATCATTTGTGCATATT 28624
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 700 TTCTTGAAATGTTCTCTAGATTGGAAGCAACCCCTCGATTCAACAGGATCAGAGCTA 759
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28623 TCTCAGAAGTGAAGATAGTTCAGATCATTTGAAGTGTCTTCTCCCAAGGATCAGCC 28564
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 760 CATATGTCCTCTCCCTTGGTGAATCCAGAGCTAAAGAACAAAGTACTGAATCTCAGAG 819
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28563 CCGACCTAGTACCGATCTCCTCGAGTTGACTCGAGAGATATAG-AGAACTTAGGGAG 28505
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 820 GATCCGAACGCAATACGGTGAACGAGGTTTCAATGAGCTGTGGATAGTATGACAT 879
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 28504 CAAGCCAAGAGGGAGTCAAGTTCGCTCTCATCATGAGCTTCACTTGTCAACAGCTTGGTT 28445
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids. 1 (bases 1 to 136047) Lin.X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M. Arabidopsis thaliana 'IGF' BAC 'F7F7' genomic sequence near marker 'mi72'

Unpublished 2 (bases 1 to 136047) Town,C.D. and Kaul,S. Direct Submission Submitted (01-JUN-2000) The Institute for Genomic Research, 9712 Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org On Jan 19, 2001 this sequence version replaced gi:12280790.

* NOTE: This is a 'working draft' sequence. It currently consists of 4 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence as soon as it is available and the accession number will be preserved.

1 100163: contig of 100163 bp in length 100164 100213: gap of unknown length 100214 112623: contig of 12410 bp in length 112624 112673: gap of unknown length 112674 116336: contig of 3963 bp in length 116337 116886: gap of unknown length 116887 136047: contig of 19361 bp in length.

Location/Qualifiers 1..136047 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="I" /clone="IGF-F7F7"

BASE COUNT 46028 a 22872 c 22213 g 44783 t 151 others ORIGIN

Query Match 3.7%; Score 63; DB 74; Length 136047; Best Local Similarity 45.1%; Pred. No. 3.2e-05; Matches 436; Conservative 0; Mismatches 520; Indels 11; Gaps 5;

QY 460 CGCTCCGTAGCGTGAAGTAAACCGTTTTCCTAACCGTCAGTACCGTGGCTCTGACGG 519 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| Db 28857 CGGTTCTTTTGGCAAGTCACTTTGTTCCCAACCAAGGGTTTGTATTGGTATCGCTT 28798 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 520 CACATCATTAATTCAGATGCTTAAAGTTTGTATGTTTCATCAATGCTTTGGGCTATTA 579 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| Db 28797 CTCATCATCTGTTATGGACGGGAACAGTGGTTAGGTTTATAAATCATGGCTCACA 28738 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 580 TTAAACAAATTTGGGAAGACGCGGACTTTGTTCCCGGAATCTTCTTCCATCTTTCGATA 639 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| Db 28737 TTTGTAACATGCA-----GCCATGGATTACCGGAGGATTTGACTCCGGTTTAGATC 28684 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 640 GATCGATAATCAAGATCTGTATGGCTAGAGGAACATTTTGAACGAATCAAGATG 699 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| Db 28683 GTACGGTTATCAATGTCCCTGCAAGTCTTGATGCAAAAATCATTTGTGCATATT 28624 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 700 TTCTTGAAATGTTCTCTAGATTGGAAGCAACCCCTCGATTCAACAGGATCAGAGCTA 759 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| Db 28623 TCTCAGAAGTGAAGATAGTTCAGATCATTTGAAGTGTCTTCTCCCAAGGATCAGCC 28564 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 760 CATATGTCCTCTCCCTTGGTGAATCCAGAGCTAAAGAACAAAGTACTGAATCTCAGAG 819 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| Db 28563 CCGACCTAGTACCGATCTCCTCGAGTTGACTCGAGAGATATAG-AGAACTTAGGGAG 28505 ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 880 GCATGGTCAAAATGACGCTGCTATCAGAGGAATCATCGAAGCAACGACGAAATGAGC 939
 Db 28444 GCCAACGCGTATCTTTGACATGCTTGTAAAGCGCTGGAGGA--GATGAGACAGAC 28387
 Qy 940 TCGAGTACTTCAGTTTACAGGGATCCGAGGACTTCTGACGCCGCCGCTGCGGCTA 999
 Db 28386 CGGTCCGTTTATGCTGCTGATTTTCAGGAACCGGTTAGATCCACCGGTTCTGAGA 28327
 Qy 1000 ACTACTTTGGCAACTGCT-TCCGTCATCGCTGTCGAAAGCAACACATTAAGAGTTAGTT 1058
 Db 28336 TGTATTTCCGGAACGCTGTTTCCGATAGGTGCTTTGGTTACAAGCGCAATGTTT 28267
 Qy 1059 GGGGATAAGGCTTCTTTGTCAGTTTCAGCTATTGAGAGAGCCATTGAAAGAGGTTG 1118
 Db 28266 TGGGAGAGATGGGTTTCTTAATATGGTTCGAAATCTTAAGTATTCGGTTAGAAGTATTG 28207
 Qy 1119 CACAAAGAAAGGGTTTCTTTCAGATGCAAAATCTGGTTATTCGGAATCTTAATGGAATC 1178
 Db 28206 GTTTACGAAAGTGGAAACAATTTGCGAATTTCTACATTAATGGAACAAGAGTGTGAAC 28147
 Qy 1179 CTTCAAAAGATTTCTCGGATACCGGATCGCTAAGTTTCGATTGCTATGGTCTAGAT 1238
 Db 28146 CAGGTACACAGA-TTGGATTCATTCGCGGTTCCGAACCACTTTGGGCTATACGGGTACAT 28088
 Qy 1239 TTTGGATGGGAAAGCCCTGCAAAATTTGACATTACCTCTGTTGATTATGCAGAATTTGATT 1298
 Db 28087 TTCGGTGGGAAACCATGTAACAGTGCAGATTGCTCCATTGACCGAAACAGGCGTTT 28028
 Qy 1299 TATGTGATTCAGTCAGGATTTTGAAGAGGTGGAGGATTCAGTATCATTCCTTAAG 1358
 Db 28027 TCGATGTCGGAGAGAGGAGTCAACCCGAGGTGTCGAGATCGGCTGTTGTTGAAGAAA 27968
 Qy 1359 ATTCAATGATGATGCTTTGCAAAATCTTTGAAGAGGCTTTTGTCTCTTTCATAGTCT 1418
 Db 27967 TGTGATGCGACATATTATTATTCTGTTTCAAAATGCTTTAGAAAATGCTTTTAGGGT 27908
 Qy 1419 CTTTAAT 1425
 Db 27907 CTTTGTAT 27901
 RESULT 15
 AB028618 85690 bp DNA PLN 27-DEC-2000
 LOCUS Abidopsis thaliana genomic DNA, chromosome 3, p1 clone:MOD1.
 DEFINITION Abidopsis thaliana genomic DNA, chromosome 3, p1 clone:MOD1.
 ACCESSION AB028618 BA00014
 VERSION AB028618.1 GI:5041971
 KEYWORDS
 SOURCE
 ORGANISM
 Abidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
 clone:MOD1.
 Abidopsis thaliana
 Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
 Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
 Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)
 Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
 Structural analysis of Arabidopsis thaliana chromosome 3. II.
 Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
 and BAC clones
 DNA Res. 7 (3), 217-221 (2000)
 2 (bases 1 to 85690)
 Kaneko,T., Katoh,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
 Direct Submission
 Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research
 Institute, Department of Plant Gene Research; 1532-3, Yana,
 Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
 Tel:81-438-52-3935, Fax:81-438-52-3934)
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/sgd_graph.cgi?c=MOD1
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant

protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Grail
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Graill-1.3),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://gremlin.zool.iastate.edu/cgi-bin/sp.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is T13J10 and the 3' clone is T26G12.
 Location/Qualifiers
 1. 85690
 /organism="Arabidopsis thaliana"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="3"
 /clone="MOD1"
 /clone_lib="Mitsui P1"
 2123. 8637
 /note="contains similarity to TNP2-like transposon
 protein
 gene_id:MOD1.1"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 join(10671..10844,10921..11208)
 /note="contains similarity to unknown protein
 emb|CAA16722.1
 gene_id:MOD1.2"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 15142..16497
 /note="gene_id:MOD1.3"
 /codon_start=1
 /evidence=not_experimental
 /product="anthocyanin 5-aromatic
 acyltransferase/benzoyltransferase-like protein"
 /protein_id="BAB02518.1"
 /db_xref="GI:11994477"
 /translation="MTLHVETARVTPTYSVINSANLHKLPLTFEFLPWLFPQVKR
 VFFELTSTRDHFHSIILPKLDSLSLILRNLYPLTGHITWEPNEPKPSLIVSENG
 VLVTIAESDAADFSLHSGYQRPUSLHALVPKLPVSDSATAFSIQITLFPNQGSFG
 VAAHVAIDGKTSTFIKAWAQICKQLQSPENLTPSYDRSLIKTYPTLDERKIELV
 RSLKEDQTNIRSLTSLPSSKIGDDVLATLVLSRADLERLREQVKNVSPSLHSTFVI
 AYAVATCFVKARGNKDRSVSLLEFGDFDRDLDPKLPQTYFGNCMLPVGYNRKAEE
 FMEKGPVTAAEIISDLVKGISLRSKTIETIADTFVEGFSQSWSTQFCTIAGSTBLGVY
 ADFGWRGPRVKVDIVSDQGEATMAERDESQGVGCMCLKKTEMDSVSVFFNNGLH
 S"
 complement(18639..19994)
 /note="gene_id:MOD1.4"
 /codon_start=1
 /evidence=not_experimental
 /product="anthocyanin 5-aromatic
 acyltransferase/benzoyltransferase-like protein"
 /protein_id="BAB02519.1"
 /db_xref="GI:11994478"
 /translation="MALNVIKISRVLSVNSVEPLVPLTFFDLWLKLNPIERVTFY
 KLFSSRDSFSSITLPLEOSLIVLSHFLPLSGHLAKNPQDPKPHIVIRPKDPLVSLT
 VVSEADFSYISSKELRELETLPLPELOVSSDSALSLSQITLFPNQGSIGITTVH
 HVMWDGKTASKFHSWAHKEHGTTPQDFDLVLDRTVINVPAGLSQKIFQLSSYIS
 EEKDYARTLTPPAKEIDNDVVRVLTLEVDIEKLEKRAKNESRDLHLSTFVFL
 AYVLTGMVKSCGGDANRPFVMAADFNRLDPEVPLTYFGNCVLPIDFNGKATFVL
 KDGKYNVGVILSDSVRLGSLRNIESIWEVYEDGCTKNMKLDQNTQVTVTGSNOFCIYS
 GFGWGRPVKTDVMSLYKNNEFSMSARDEIGGLEIGISLKKCNENFVLSLFTSDFDIY
 K"
 complement(join(27960..28438,28676..28739))
 /note="gene_id:MOD1.5"

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:26 ; Search time 445.49 seconds
(without alignments)
2262.188 Million cell updates/sec

Title: US-08-894-356C-3

Perfect score: 1605

Sequence: 1 TGTCGACGAATCCATTCA.....TTAAAAAAAAAAAAAAAAAAAA 1605

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

```
1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	1601.8	99.8	1605	17	AAT37310 Aromatic acyl tran
2	512.4	31.9	1815	21	AAC48422 Arabidopsis thalia
3	493.6	30.8	1813	21	AAC37537 Arabidopsis thalia
4	378.8	23.6	1359	21	AAC48421 Arabidopsis thalia
5	370.8	23.1	1363	21	AAC40007 Arabidopsis thalia
6	361.6	22.5	2412	21	AAC48416 Arabidopsis thalia
7	140	8.7	1364	21	AAC48415 Arabidopsis thalia
8	136.6	8.5	462	21	AAC48415 Arabidopsis thalia
9	132	8.2	1363	21	AAC39553 Arabidopsis thalia
10	106.2	6.6	1520	21	AAC42254 Arabidopsis thalia
11	81	5.0	936	22	AAF58252 Oligonucleotide D1

12	81	5.0	936	22	AAF58254 Oligonucleotide D1
13	81	5.0	936	22	AAF58257 Oligonucleotide D1
14	81	5.0	936	22	AAF58259 Oligonucleotide D2
15	81	5.0	936	22	AAF58262 Oligonucleotide D2
16	81	5.0	936	22	AAF58255 Oligonucleotide D1
17	79.2	4.9	936	22	AAF58252 Oligonucleotide D1
18	79.2	4.9	936	22	AAF58254 Oligonucleotide D1
19	79.2	4.9	936	22	AAF58257 Oligonucleotide D1
20	79.2	4.9	936	22	AAF58259 Oligonucleotide D2
21	79.2	4.9	936	22	AAF58262 Oligonucleotide D2
22	79.2	4.9	936	22	AAF58255 Oligonucleotide D1
23	67	4.2	1477	21	AAC39128 Arabidopsis thalia
24	58.6	3.7	1648	21	AAC64785 Arabidopsis thalia
25	57.4	3.6	3399	17	AAT05868 Chicken leucocytos
26	54.8	3.4	1429	21	AAC33965 Arabidopsis thalia
27	53.6	3.3	1498	21	AAC51224 Arabidopsis thalia
28	53.6	3.3	1500	21	AAC36258 Arabidopsis thalia
29	53.2	3.3	1427	21	AAC49764 Arabidopsis thalia
30	53	3.3	244	22	AAF58238 Oligonucleotide D1
31	52.8	3.3	244	22	AAF58238 Oligonucleotide D1
32	51.2	3.2	3579	21	AA70099 Plasmodium falcipa
33	50	3.1	2427	11	AAQ04107 Human pro-urokinas
34	49.4	3.1	2377	13	AAQ20360 Human pro-urokinas
35	47	2.9	926	21	AAA61262 Homo sapiens CO122
36	46.8	2.9	2214	19	AAV40524 Plasmodium falcipa
37	46.4	2.9	1299	21	AAA70111 Homo sapiens CO122
38	45.4	2.8	2814	21	AAA88890 Mouse Wnt-3a gene
39	45.4	2.8	2814	21	AAZ47790 Murine Wnt-3a enco
40	45	2.8	1043	13	AAQ21651 Macromolecular sys
41	44.8	2.8	1474	18	AAT90174 Oil seed rape cyst
42	44.6	2.8	729	20	AAK61443 DNA encoding a hum
43	44.6	2.8	1672	22	AAF92066 Human PRO1063 cdna
44	44.6	2.8	1701	21	AAZ64982 Membrane-bound pro
45	44.6	2.8	1701	22	AAF44128 Human PRO4063 (UNQ

ALIGNMENTS

RESULT 1

AAT37310

ID AAT37310 standard; cdna to mRNA; 1605 BP.

AC AAT37310;

AC AAT37310;

DT 06-FEB-1997 (first entry)

XX

DE Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation; anthocyanin pigment;

KW plants; acylation; colour; tone; colouration; colour change;

KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;

KW Scenecio cruentus; Lavandula angustifolia; ds.

XX

OS Petunia hybrida (Clone pPAT48).

XX

EH Key Location/Qualifiers

FT CDS 67..1413

FT /*tag= a

FT /product= Aromatic acyl transferase.

XX

PN WO9625500-A1.

XX

PD 22-AUG-1996.

XX

PF 16-FEB-1996; 96WO-JP00348.

XX

PR 30-JAN-1996; 96JP-0046534.

XX

PR 17-FEB-1995; 95JP-0067159.

XX

PA 29-JUN-1995; 95JP-0196915.

XX

XX (SUNR) SUNTORY LTD.

RESULT	2	
AAC48422		
ID	AAC48422	standard; DNA; 1815 BP.
XX		
XX		
AC	AAC48422;	
XX		
DT	18-OCT-2000	(first entry)
XX		
XX		
DE	Arabidopsis	thaliana DNA fragment
XX		
XX		SEQ ID NO: 57422.

KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0128845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 23-APR-1999; 99US-0130891.

PR 28-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 30-APR-1999; 99US-0132407.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 06-MAY-1999; 99US-0132487.

PR 07-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR 23-JUN-1999; 99US-0140353.

PR 23-JUN-1999; 99US-0140354.

PR 24-JUN-1999; 99US-0140695.

PR 28-JUN-1999; 99US-0140823.

PR 29-JUN-1999; 99US-0140991.

PR 30-JUN-1999; 99US-0141287.

PR 01-JUL-1999; 99US-0141842.

PR 01-JUL-1999; 99US-0142154.

PR 02-JUL-1999; 99US-0142055.

PR 06-JUL-1999; 99US-0142390.

PR 08-JUL-1999; 99US-0142803.

PR 09-JUL-1999; 99US-0142920.

PR 12-JUL-1999; 99US-0142977.

PR 13-JUL-1999; 99US-0143542.

PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.

PR 16-JUL-1999; 99US-0144086.

PR 19-JUL-1999; 99US-0144325.

PR 19-JUL-1999; 99US-0144331.

PR 19-JUL-1999; 99US-0144332.

PR 19-JUL-1999; 99US-0144333.

PR 19-JUL-1999; 99US-0144334.

PR 19-JUL-1999; 99US-0144335.

PR 20-JUL-1999; 99US-0144352.

PR 20-JUL-1999; 99US-0144632.

PR 20-JUL-1999; 99US-0144884.

PR 21-JUL-1999; 99US-0144814.

PR 21-JUL-1999; 99US-0145086.

PR 21-JUL-1999; 99US-0145088.

PR 22-JUL-1999; 99US-0145085.

PR 22-JUL-1999; 99US-0145087.

PR 22-JUL-1999; 99US-0145089.

PR 22-JUL-1999; 99US-0145192.

PR 23-JUL-1999; 99US-0145145.

PR 23-JUL-1999; 99US-0145218.

PR 23-JUL-1999; 99US-0145224.

PR 26-JUL-1999; 99US-0145276.

PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.

PR 27-JUL-1999; 99US-0145919.

PR 28-JUL-1999; 99US-0145951.

PR 02-AUG-1999; 99US-0146386.

PR 02-AUG-1999; 99US-0146388.

PR 02-AUG-1999; 99US-0146389.

PR 03-AUG-1999; 99US-0147038.

PR 04-AUG-1999; 99US-0147204.

PR 04-AUG-1999; 99US-0147302.

PR 05-AUG-1999; 99US-0147192.

PR 05-AUG-1999; 99US-0147260.

PR 06-AUG-1999; 99US-0147303.

PR 06-AUG-1999; 99US-0147456.

PR 09-AUG-1999; 99US-0147493.

PR 09-AUG-1999; 99US-0147915.

PR 10-AUG-1999; 99US-0148171.

PR 11-AUG-1999; 99US-0148319.

PR 12-AUG-1999; 99US-0148341.

PR 13-AUG-1999; 99US-0148565.

PR 13-AUG-1999; 99US-0148684.

PR 16-AUG-1999; 99US-0149368.

PR 17-AUG-1999; 99US-0149175.

PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.

PR 20-AUG-1999; 99US-0149723.

PR 20-AUG-1999; 99US-0149929.

PR 23-AUG-1999; 99US-0149902.

PR 23-AUG-1999; 99US-0149930.

PR 25-AUG-1999; 99US-0150556.

PR 26-AUG-1999; 99US-0150834.

PR 27-AUG-1999; 99US-0151035.

PR 27-AUG-1999; 99US-0151036.

PR 27-AUG-1999; 99US-0151030.

PR 30-AUG-1999; 99US-0151303.

PR 31-AUG-1999; 99US-0151438.

ID AAC37537 standard; DNA; 1813 BP.
XX AC AAC37537;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 17744.
XX KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PE 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148555.
PR 13-AUG-1999; 99US-0148634.
PR 16-AUG-1999; 99US-0149358.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.

QY 1375 TTGGAGAAAGATAAAGAGTT 1394
| | | | |
Db 1435 cttgtgaagagcaagagtt 1454

RESULT 4
AAC48421
ID AAC48421 standard; DNA; 1359 BP.
XX
AC AAC48421;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 57418.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0132863.
PR 14-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139464.
PR 18-JUN-1999; 99US-0139465.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139819.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140891.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 02-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142603.
PR 09-JUL-1999; 99US-0142620.
PR 12-JUL-1999; 99US-0142677.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144684.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 06-AUG-1999; 99US-0147250.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148555.

PD	06-SEP-2000.	PR	06-JUL-1999;	99US-0142390.
XX		PR	08-JUL-1999;	99US-0142803.
PF		PR	09-JUL-1999;	99US-0142820.
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0142877.
PR	25-FEB-1999;	PR	13-JUL-1999;	99US-0143442.
PR	05-MAR-1999;	PR	14-JUL-1999;	99US-0144005.
PR	09-MAR-1999;	PR	15-JUL-1999;	99US-0144085.
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086.
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325.
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331.
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332.
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333.
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334.
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335.
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352.
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632.
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884.
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144814.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086.
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088.
PR	04-MAY-1999;	PR	22-JUL-1999;	99US-0145085.
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089.
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192.
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145.
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218.
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224.
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913.
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918.
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919.
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951.
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386.
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388.
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389.
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038.
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204.
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302.
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192.
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260.
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303.
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416.
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493.
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935.
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171.
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319.
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341.
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565.
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684.
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368.
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175.
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723.
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902.
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930.
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566.
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151065.
PR	18-JUN-1999;	PR	27-AUG-1999;	99US-0151066.
PR	21-JUN-1999;	PR	27-AUG-1999;	99US-0151080.
PR	22-JUN-1999;	PR	30-AUG-1999;	99US-0151303.
PR	23-JUN-1999;	PR	31-AUG-1999;	99US-0151338.
PR	23-JUN-1999;	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	PR	07-SEP-1999;	99US-0152353.
PR	24-JUN-1999;	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	PR	22-SEP-1999;	99US-0155139.

PR	23-SEP-1999;	99US-0155486.	
PR	24-SEP-1999;	99US-0155659.	
PR	28-SEP-1999;	99US-0156458.	
PR	29-SEP-1999;	99US-0156596.	
PR	04-OCT-1999;	99US-0157117.	
PR	05-OCT-1999;	99US-0157753.	
PR	06-OCT-1999;	99US-0157865.	
PR	07-OCT-1999;	99US-0158029.	
PR	08-OCT-1999;	99US-0158232.	
PR	12-OCT-1999;	99US-0158369.	
PR	13-OCT-1999;	99US-0159293.	
PR	13-OCT-1999;	99US-0159294.	
PR	13-OCT-1999;	99US-0159295.	
PR	14-OCT-1999;	99US-0159329.	
PR	14-OCT-1999;	99US-0159330.	
PR	14-OCT-1999;	99US-0159331.	
PR	14-OCT-1999;	99US-0159637.	
PR	14-OCT-1999;	99US-0159638.	
PR	18-OCT-1999;	99US-0159584.	
PR	21-OCT-1999;	99US-0160741.	
PR	21-OCT-1999;	99US-0160767.	
PR	21-OCT-1999;	99US-0160768.	
PR	21-OCT-1999;	99US-0160770.	
PR	21-OCT-1999;	99US-0160814.	
PR	22-OCT-1999;	99US-0160815.	
PR	22-OCT-1999;	99US-0160980.	
PR	22-OCT-1999;	99US-0160981.	
PR	22-OCT-1999;	99US-0160989.	
PR	25-OCT-1999;	99US-0161404.	
PR	25-OCT-1999;	99US-0161405.	
PR	25-OCT-1999;	99US-0161406.	
PR	26-OCT-1999;	99US-0161359.	
PR	26-OCT-1999;	99US-0161360.	
PR	26-OCT-1999;	99US-0161361.	
PR	28-OCT-1999;	99US-0161920.	
PR	28-OCT-1999;	99US-0161992.	
PR	28-OCT-1999;	99US-0161993.	
PR	29-OCT-1999;	99US-0162142.	
Query Match 23.1%; Score 370.8; DB 21; Length 1363;			
Best Local Similarity 65.3%; Pred. No. 4.6e-81;			
Matches 581; Conservative 0; Mismatches 297; Indels 12; Gaps 2;			
Qy	514	CAGCTCAACAAAGCTCAAGGACGGGCTCACCATGGGATTAGCAATTAACCATGCTGTGCTG 573	
Db	117	cggtgaccaaagctaaagatggcttgcaatggcctagcttccaatcacgcagctcta 176	
Qy	574	CATGGTACTTCGACGTGGCACTTTATGACCTCGTGGTCCGAGCTTTGCTGTGGTCCACC 633	
Db	177	gacggacatccacgtggcattcatgagctcatggccgagatctgctggagcccaa 236	
Qy	634	TCAATTTCTGCCACCATTCCTTGAAGACCAAGGCTCGTAACACTCGAGTCAAGCTC 693	
Db	237	tccatctcgcccaaaccttctcttgaccgatcggaagcgcgtgacacgcgctgaagctc 296	
Qy	694	AACCTCTCTCAACCATCAGATGACCGGACATGCTAAGTCAGCAACCAACGCGTATGTC 753	
Db	297	gatctactgtctccaaagaccaccaagacacctccaacgcgaagatgcgcaccaat--- 353	
Qy	754	CCGGCCAAAGTAGACCAACCTCTTCGGGAAAGAGATTATCAAGTTCTCCGAGTTAGCAATT 813	
Db	354	ccgacggtggagccgcgcagctagtagaagaatcttcaggttctcggaacttgctgctc 413	
Qy	814	GACAAATCAAGTCAA-----CAGTCAATGCCAACTCAGGAGAGACCCATTCTCC 864	
Db	414	cacacgatcaagtcacaagtgctaatctcagtaatcccatcggaagcgttcaaacacattctcc 473	
Qy	865	ACATTCGAATCACTCTCCGCACACGCTGTGGCTAGCCGCTACACGTCGCGCCCAACTCAAG 924	
Db	474	acttccaataactcacatcacatctgpgccacgtcacttagcgcgtggagctcaaa 533	
Qy	925	CCCAGGACTACACTGTGTACACTGTGTGCTGATTCGACGAGAAAGGTTGATCCTCCA 984	

Db	534	ccggagacataaactatctctaccgcttcttcaccgcttcctcgccgactgtccgcccgcgttgatcctcca 593	
Qy	985	ATGCCAGAAAGTTACTTCGGCAACCTAATTACAGGCAATTTTCACAGTACGCGCGCAGGT 1044	
Db	594	atccggaggaataatttcgaaacttgattcaagccattctccacggaacagcgcgggg 653	
Qy	1045	TTGTTACTAGCAAGCCGATYCGAGTTGCTGCTGGTGGATGATACAAACGATCGTGAAG 1104	
Db	654	ctactagcagcgacgcgaccgagctcggagcttcagtgatccagaaagcgcgcgca 713	
Qy	1105	CATGACCTAAGCCATTGTATGAAGAAACAGAGGTGGGAGAGCAACCCGAAGATCTTT 1164	
Db	714	cacgacgcgagtgcttaacgacgcgttaacgacgagtcgggagaaatccacgaaagatttt 773	
Qy	1165	CAGTACAAAGATGCTGGAGTGAAGTGTGCTGCTGGAAGTTCGCCAAGGTTCAAGGTT 1224	
Db	774	caattcaaacgacgcggagtgaaattgtgcggtgggaagctcaccgaggttcctgctt 833	
Qy	1225	TACGACGTGATTTTGGATGGGAAAGCCAGAGAGTGTGAGGAGTGTTCGAACAATAGG 1284	
Db	834	tacgaggtggatttcggttttgaaaccggaacgcgttaggagcgggtcaataaacgcg 893	
Qy	1285	TTTCATGGAATGCTGTTATTTGTACCAAGGCACAAATGGAGAAAGCAAGCATTTGATGGAG 1344	
Db	894	tttaatgggatgatgtatttgtaaccaaggtaaacgacgagcataagattattgatgtgag 953	
Qy	1345	ATTAGTTTGGAAAGCAATGCTATGGAGAGGTTGGAGAAAGATAAAGAGTT 1394	
Db	954	ataactttgaagcttctgttatggagaagcttctgaagagcaaaagatt 1003	
RESULT 6			
ID	AAC48416		
XX	AAC48416 standard; DNA; 2412 BP.		
AC	AAC48416;		
XX			
DT	18-OCT-2000 (first entry)		
XX			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 57400.		
XX			
KW	Hybridisation assay; genetic mapping; gene expression control;		
KW	protein identification; signal transduction pathway;		
KW	metabolic pathway; promoter; termination sequence; ss.		
XX			
OS	Arabidopsis thaliana.		
XX			
PN	EP1033405-A2.		
XX			
PD	06-SEP-2000.		
XX			
PF	25-FEB-2000; 2000EP-0301439.		
XX			
PR	25-FEB-1999; 99US-0121825.		
PR	05-MAR-1999; 99US-0123180.		
PR	09-MAR-1999; 99US-0123548.		
PR	23-MAR-1999; 99US-0125788.		
PR	25-MAR-1999; 99US-0126264.		
PR	29-MAR-1999; 99US-0126785.		
PR	01-APR-1999; 99US-0127462.		
PR	06-APR-1999; 99US-0128234.		
PR	08-APR-1999; 99US-0128714.		
PR	16-APR-1999; 99US-0129845.		
PR	19-APR-1999; 99US-0130077.		
PR	21-APR-1999; 99US-0130449.		
PR	23-APR-1999; 99US-0130510.		
PR	28-APR-1999; 99US-0130891.		
PR	28-APR-1999; 99US-0131449.		
PR	30-APR-1999; 99US-0132048.		
PR	30-APR-1999; 99US-0132407.		
PR	04-MAY-1999; 99US-0132484.		
PR	05-MAY-1999; 99US-0132485.		
PR	06-MAY-1999; 99US-0132486.		

PR 06-MAY-1999;	990S-0132487.	PR 22-JUL-1999;	990S-0145192.
PR 07-MAY-1999;	990S-0132863.	PR 23-JUL-1999;	990S-0145145.
PR 11-MAY-1999;	990S-0134256.	PR 23-JUL-1999;	990S-0145218.
PR 14-MAY-1999;	990S-0134218.	PR 23-JUL-1999;	990S-0145224.
PR 14-MAY-1999;	990S-0134219.	PR 26-JUL-1999;	990S-0145276.
PR 14-MAY-1999;	990S-0134221.	PR 27-JUL-1999;	990S-0145913.
PR 14-MAY-1999;	990S-0134370.	PR 27-JUL-1999;	990S-0145918.
PR 18-MAY-1999;	990S-0134768.	PR 27-JUL-1999;	990S-0145919.
PR 19-MAY-1999;	990S-0134941.	PR 28-JUL-1999;	990S-0145951.
PR 20-MAY-1999;	990S-0135124.	PR 02-AUG-1999;	990S-0146386.
PR 21-MAY-1999;	990S-0135353.	PR 02-AUG-1999;	990S-0146388.
PR 24-MAY-1999;	990S-0135629.	PR 02-AUG-1999;	990S-0146389.
PR 25-MAY-1999;	990S-0136021.	PR 03-AUG-1999;	990S-0147033.
PR 27-MAY-1999;	990S-0136392.	PR 04-AUG-1999;	990S-0147204.
PR 28-MAY-1999;	990S-0136782.	PR 04-AUG-1999;	990S-0147302.
PR 01-JUN-1999;	990S-0137222.	PR 05-AUG-1999;	990S-0147192.
PR 03-JUN-1999;	990S-0137528.	PR 05-AUG-1999;	990S-0147260.
PR 04-JUN-1999;	990S-0137502.	PR 06-AUG-1999;	990S-0147303.
PR 07-JUN-1999;	990S-0137724.	PR 06-AUG-1999;	990S-0147416.
PR 08-JUN-1999;	990S-0138094.	PR 09-AUG-1999;	990S-0147493.
PR 10-JUN-1999;	990S-0138540.	PR 09-AUG-1999;	990S-0147935.
PR 10-JUN-1999;	990S-0138847.	PR 10-AUG-1999;	990S-0148171.
PR 14-JUN-1999;	990S-0139119.	PR 11-AUG-1999;	990S-0148319.
PR 16-JUN-1999;	990S-0139452.	PR 12-AUG-1999;	990S-0148341.
PR 16-JUN-1999;	990S-0139453.	PR 13-AUG-1999;	990S-0148565.
PR 17-JUN-1999;	990S-0139492.	PR 13-AUG-1999;	990S-0148684.
PR 18-JUN-1999;	990S-0139454.	PR 16-AUG-1999;	990S-0149368.
PR 18-JUN-1999;	990S-0139455.	PR 17-AUG-1999;	990S-0149175.
PR 18-JUN-1999;	990S-0139456.	PR 18-AUG-1999;	990S-0149426.
PR 18-JUN-1999;	990S-0139457.	PR 20-AUG-1999;	990S-0149722.
PR 18-JUN-1999;	990S-0139458.	PR 20-AUG-1999;	990S-0149723.
PR 18-JUN-1999;	990S-0139459.	PR 20-AUG-1999;	990S-0149929.
PR 18-JUN-1999;	990S-0139460.	PR 23-AUG-1999;	990S-0149902.
PR 18-JUN-1999;	990S-0139461.	PR 23-AUG-1999;	990S-0149930.
PR 18-JUN-1999;	990S-0139462.	PR 25-AUG-1999;	990S-0150566.
PR 18-JUN-1999;	990S-0139463.	PR 26-AUG-1999;	990S-0150884.
PR 18-JUN-1999;	990S-0139750.	PR 27-AUG-1999;	990S-0151065.
PR 18-JUN-1999;	990S-0139763.	PR 27-AUG-1999;	990S-0151066.
PR 21-JUN-1999;	990S-0139817.	PR 27-AUG-1999;	990S-0151080.
PR 22-JUN-1999;	990S-0139899.	PR 30-AUG-1999;	990S-0151303.
PR 23-JUN-1999;	990S-0140353.	PR 31-AUG-1999;	990S-0151438.
PR 23-JUN-1999;	990S-0140354.	PR 01-SEP-1999;	990S-0151910.
PR 24-JUN-1999;	990S-0140695.	PR 07-SEP-1999;	990S-0152363.
PR 28-JUN-1999;	990S-0140823.	PR 10-SEP-1999;	990S-0153070.
PR 29-JUN-1999;	990S-0140991.	PR 13-SEP-1999;	990S-0153758.
PR 30-JUN-1999;	990S-0141287.	PR 15-SEP-1999;	990S-0154018.
PR 01-JUL-1999;	990S-0141842.	PR 16-SEP-1999;	990S-0154039.
PR 01-JUL-1999;	990S-0142154.	PR 20-SEP-1999;	990S-0154779.
PR 02-JUL-1999;	990S-0142055.	PR 22-SEP-1999;	990S-0155139.
PR 06-JUL-1999;	990S-0142390.	PR 23-SEP-1999;	990S-0155486.
PR 08-JUL-1999;	990S-0142803.	PR 24-SEP-1999;	990S-0155659.
PR 09-JUL-1999;	990S-0142920.	PR 28-SEP-1999;	990S-0156458.
PR 12-JUL-1999;	990S-0142977.	PR 29-SEP-1999;	990S-0156596.
PR 13-JUL-1999;	990S-0143542.	PR 04-OCT-1999;	990S-0157117.
PR 14-JUL-1999;	990S-0143624.	PR 05-OCT-1999;	990S-0157753.
PR 15-JUL-1999;	990S-0144005.	PR 06-OCT-1999;	990S-0157865.
PR 16-JUL-1999;	990S-0144085.	PR 07-OCT-1999;	990S-0158029.
PR 16-JUL-1999;	990S-0144086.	PR 08-OCT-1999;	990S-0158232.
PR 19-JUL-1999;	990S-0144325.	PR 12-OCT-1999;	990S-0158369.
PR 19-JUL-1999;	990S-0144331.	PR 13-OCT-1999;	990S-0159293.
PR 19-JUL-1999;	990S-0144332.	PR 13-OCT-1999;	990S-0159294.
PR 19-JUL-1999;	990S-0144333.	PR 13-OCT-1999;	990S-0159295.
PR 19-JUL-1999;	990S-0144334.	PR 14-OCT-1999;	990S-0159329.
PR 19-JUL-1999;	990S-0144335.	PR 14-OCT-1999;	990S-0159330.
PR 20-JUL-1999;	990S-0144352.	PR 14-OCT-1999;	990S-0159331.
PR 20-JUL-1999;	990S-0144632.	PR 14-OCT-1999;	990S-0159637.
PR 20-JUL-1999;	990S-0144884.	PR 14-OCT-1999;	990S-0159638.
PR 21-JUL-1999;	990S-0144814.	PR 18-OCT-1999;	990S-0159534.
PR 21-JUL-1999;	990S-0145086.	PR 21-OCT-1999;	990S-0160711.
PR 21-JUL-1999;	990S-0145088.	PR 21-OCT-1999;	990S-0160757.
PR 22-JUL-1999;	990S-0145085.	PR 21-OCT-1999;	990S-0160758.
PR 22-JUL-1999;	990S-0145087.	PR 21-OCT-1999;	990S-0160770.
PR 22-JUL-1999;	990S-0145089.	PR 21-OCT-1999;	990S-0160814.

PR	16-JUN-1999;	99US-0139453.
PR	17-JUN-1999;	99US-0139492.
PR	18-JUN-1999;	99US-0139454.
PR	18-JUN-1999;	99US-0139455.
PR	18-JUN-1999;	99US-0139456.
PR	18-JUN-1999;	99US-0139457.
PR	18-JUN-1999;	99US-0139458.
PR	18-JUN-1999;	99US-0139459.
PR	18-JUN-1999;	99US-0139460.
PR	18-JUN-1999;	99US-0139461.
PR	18-JUN-1999;	99US-0139462.
PR	18-JUN-1999;	99US-0139463.
PR	18-JUN-1999;	99US-0139750.
PR	21-JUN-1999;	99US-0139763.
PR	21-JUN-1999;	99US-0139817.
PR	22-JUN-1999;	99US-0139899.
PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140695.
PR	28-JUN-1999;	99US-0140823.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145192.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0145951.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159253.
PR	13-OCT-1999;	99US-0159254.
PR	13-OCT-1999;	99US-0159255.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.
Query Match		
Best Local Similarity 8.7%; Score 140; DB 21; Length 1364;		
Pred. No. 1.1e-24;		
Matches 469; Conservative 0; Mismatches 500; Indels 9; Gaps 2;		
QY	423	TGATGAAGAGGACCACCAAAATTCGAGACTTGTGATTCCTTGTAAATAAAATCTTGAAATTT 482
Db	374	tgctaagtatgtcccttcaattgttcaatctttctttgaccatcacaagctgtgaatca 433
QY	483	GGAAGGGCTTCATCGGCCCTCTTCTAGCTGTGCAGCTCACCAGCTCAAGGACGGGCTCAC 542

Db 434 tgatggacacaccatgagccctcttctctgtccagtgacagaattggtagatggaatttt 493
Qy 543 CATGGATTAGCATTTAACCATGCTGCTGCTGATGTTCTGACGTGGCACTTTATGAC 602
Db 494 cataggactgtctatgaacctcgtatgagggagcgtactgctgctggaagtctcttac 553
Qy 603 CTCGTGTCGAGCTTGTCTGTGGTCCACCTCAATTTCTGTCCCACTCAATTCCTTGAACG 662
Db 554 cgtatggtccgaatacttccaaagacaaagagacacaaacgaactgactgtgtcttaa 613
Qy 663 AACCAAGCTCGTACACTCGAGTCAAGCTCAAGCTCTCTCAACCATCAGATGCACCCGA 722
Db 614 gaattccaccctgttgaagcgtatcatcctgaaggaacacggtctcttattcagccttcc 673
Qy 723 ACATGCTAAGTCAGCAACCAACGGTGATGCCGCGCAACCTAGACCCACCTCTTCGCGA 782
Db 674 atatagtcacccg-----atgaattatcagaacctacgaattctcctcaagga 727
Qy 783 AAGAGTATTCAAGTTCTCCGAGTTAGCAATTCACAAATCAAGTCAACAGTCAATGCCAA 842
Db 728 gagaatgtctgttctctcatcagaacaccataagaatgctttaaacaagggtcaacccaaat 787
Qy 843 CTCAGGAGAGAGCCATTTCTCCATTCACATCCCAATCACTCTCCGCAACAGTGTGGCTAGCCGT 902
Db 788 atgcggaacaacctccatctctatcttccaatcgtacacgcggttatatgagagatgcat 847
Qy 903 CACAGTGCAGCGCCCAACCAACGCGGAGGACTACACTGTGTACACTGTGTTGCTGATTG 962
Db 848 aacaagggcagaggatctacctctcgtatcgagaaacaagtctgtgaggtgctgctgacaa 907
Qy 963 CAGAAAGAGGTTGATCTCCCAATGCCAAGAAAGTTACTTTCGGCAACCTAAATTCAGGCAAT 1022
Db 908 cagaggaagaatgatccacgcgttcacaaggattactctggaacactgcctctgcttt 967
Qy 1023 TTTACAGTGCAGCGGCGAGGTTGTTTACTAGCAAGCCCGATCAGTTCGCTGTGGGAT 1082
Db 968 gagaacggctgcaaaagctggtgaactattggagaatgatcttggattgctgctttgaa 1027
Qy 1083 GATACAACAAGCTCGTGAAGCATGACGCTAAGGCCATTCATGAAGAAACAAGGAGTG 1142
Db 1028 agtgcatacaagtagctgaacatacagagcgaaaggtatctcaaatgattgatcaatg 1087
Qy 1143 GGAGAGCAACCCGAAGATCTTTCA---GTACAAAGATGCTGGAGTGAACCTGTGTGCTGT 1199
Db 1088 gctaaagtctcttatattaccattatgctgctattgtaacccgatgagcgttatgat 1147
Qy 1200 TGGAGTTTCGCAAGGTTCAAGTTTACGACGTGGATTTTCGATGGGGAAGCCAGAGAG 1259
Db 1148 gggaaagctcgcgaggtttaaagtagtggttgtaattgggttgggaaagagtgac 1207
Qy 1260 TGTGAGGAGTGTTCGAACAATAGTTTGTATGGAATGCTGATTTGTACCAAGGCAAAA 1319
Db 1208 gcttagaagtggctatgcgcataagttcgatcggaagagatctcttaccctccgggaagaga 1267
Qy 1320 TGGAGGAAGAGCATGATGTGGAGATTAGTTTGGAGCAAAATGCTATGGAGAGGTTTGA 1379
Db 1268 aggggggggaagtagatttggaggtatgctgttccagatttatggagcgtttgga 1327
Qy 1380 GAAGATAAAGATTCCT 1397
Db 1328 atcagatgaagagttcat 1345

RESULT 8

AAC48415

ID AAC48415 standard; DNA; 462 BP.

XX AAC48415;

XX AC

XX AC

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 57397.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

XX 06-SEP-2000.

XX 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

PR 28-APR-1999; 99US-0130891.

PR 30-APR-1999; 99US-0131449.

PR 30-APR-1999; 99US-0132048.

PR 04-MAY-1999; 99US-0132484.

PR 05-MAY-1999; 99US-0132485.

PR 06-MAY-1999; 99US-0132486.

PR 07-MAY-1999; 99US-0132487.

PR 11-MAY-1999; 99US-0132863.

PR 11-MAY-1999; 99US-0134256.

PR 14-MAY-1999; 99US-0134218.

PR 14-MAY-1999; 99US-0134219.

PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.

PR 18-MAY-1999; 99US-0134768.

PR 19-MAY-1999; 99US-0134941.

PR 20-MAY-1999; 99US-0135124.

PR 21-MAY-1999; 99US-0135353.

PR 24-MAY-1999; 99US-0135629.

PR 25-MAY-1999; 99US-0136021.

PR 27-MAY-1999; 99US-0136392.

PR 28-MAY-1999; 99US-0136782.

PR 01-JUN-1999; 99US-0137222.

PR 03-JUN-1999; 99US-0137528.

PR 04-JUN-1999; 99US-0137502.

PR 07-JUN-1999; 99US-0137724.

PR 08-JUN-1999; 99US-0138094.

PR 10-JUN-1999; 99US-0138540.

PR 10-JUN-1999; 99US-0138847.

PR 14-JUN-1999; 99US-0139119.

PR 16-JUN-1999; 99US-0139452.

PR 16-JUN-1999; 99US-0139453.

PR 17-JUN-1999; 99US-0139492.

PR 18-JUN-1999; 99US-0139454.

PR 18-JUN-1999; 99US-0139455.

PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.

PR 18-JUN-1999; 99US-0139458.

PR 18-JUN-1999; 99US-0139459.

PR 18-JUN-1999; 99US-0139460.

PR 18-JUN-1999; 99US-0139461.

PR 18-JUN-1999; 99US-0139462.

PR 18-JUN-1999; 99US-0139463.

PR 18-JUN-1999; 99US-0139750.

PR 18-JUN-1999; 99US-0139763.

PR 21-JUN-1999; 99US-0139817.

PR 22-JUN-1999; 99US-0139899.

PR	23-JUN-1999;	99US-0140353.	PR	31-AUG-1999;	99US-0151433.
PR	23-JUN-1999;	99US-0140354.	PR	01-SEP-1999;	99US-0151930.
PR	24-JUN-1999;	99US-0140695.	PR	07-SEP-1999;	99US-0152363.
PR	28-JUN-1999;	99US-0140823.	PR	10-SEP-1999;	99US-0153070.
PR	29-JUN-1999;	99US-0140991.	PR	13-SEP-1999;	99US-0153758.
PR	30-JUN-1999;	99US-0141287.	PR	15-SEP-1999;	99US-0154018.
PR	01-JUL-1999;	99US-0141842.	PR	16-SEP-1999;	99US-0154039.
PR	01-JUL-1999;	99US-0142154.	PR	20-SEP-1999;	99US-0154779.
PR	02-JUL-1999;	99US-0142055.	PR	22-SEP-1999;	99US-0155139.
PR	06-JUL-1999;	99US-0142390.	PR	23-SEP-1999;	99US-0155486.
PR	08-JUL-1999;	99US-0142803.	PR	24-SEP-1999;	99US-0155659.
PR	09-JUL-1999;	99US-0142920.	PR	28-SEP-1999;	99US-0156458.
PR	12-JUL-1999;	99US-0142977.	PR	29-SEP-1999;	99US-0156596.
PR	13-JUL-1999;	99US-0143542.	PR	04-OCT-1999;	99US-0157117.
PR	14-JUL-1999;	99US-0143624.	PR	05-OCT-1999;	99US-0157753.
PR	15-JUL-1999;	99US-0144005.	PR	06-OCT-1999;	99US-0157865.
PR	16-JUL-1999;	99US-0144085.	PR	07-OCT-1999;	99US-0158029.
PR	16-JUL-1999;	99US-0144086.	PR	08-OCT-1999;	99US-0158232.
PR	19-JUL-1999;	99US-0144325.	PR	12-OCT-1999;	99US-0158369.
PR	19-JUL-1999;	99US-0144331.	PR	13-OCT-1999;	99US-0159293.
PR	19-JUL-1999;	99US-0144332.	PR	13-OCT-1999;	99US-0159294.
PR	19-JUL-1999;	99US-0144333.	PR	13-OCT-1999;	99US-0159295.
PR	19-JUL-1999;	99US-0144334.	PR	14-OCT-1999;	99US-0159329.
PR	19-JUL-1999;	99US-0144335.	PR	14-OCT-1999;	99US-0159330.
PR	20-JUL-1999;	99US-0144632.	PR	14-OCT-1999;	99US-0159331.
PR	20-JUL-1999;	99US-0144684.	PR	14-OCT-1999;	99US-0159637.
PR	21-JUL-1999;	99US-0144814.	PR	14-OCT-1999;	99US-0159638.
PR	21-JUL-1999;	99US-0145086.	PR	18-OCT-1999;	99US-0159584.
PR	21-JUL-1999;	99US-0145088.	PR	21-OCT-1999;	99US-0160741.
PR	21-JUL-1999;	99US-0145089.	PR	21-OCT-1999;	99US-0160767.
PR	22-JUL-1999;	99US-0145085.	PR	21-OCT-1999;	99US-0160768.
PR	22-JUL-1999;	99US-0145087.	PR	21-OCT-1999;	99US-0160770.
PR	22-JUL-1999;	99US-0145192.	PR	21-OCT-1999;	99US-0160814.
PR	23-JUL-1999;	99US-0145145.	PR	21-OCT-1999;	99US-0160815.
PR	23-JUL-1999;	99US-0145218.	PR	22-OCT-1999;	99US-0160980.
PR	23-JUL-1999;	99US-0145224.	PR	22-OCT-1999;	99US-0160981.
PR	26-JUL-1999;	99US-0145276.	PR	22-OCT-1999;	99US-0160989.
PR	27-JUL-1999;	99US-0145913.	PR	25-OCT-1999;	99US-0161405.
PR	27-JUL-1999;	99US-0145918.	PR	25-OCT-1999;	99US-0161406.
PR	27-JUL-1999;	99US-0145919.	PR	25-OCT-1999;	99US-0161359.
PR	28-JUL-1999;	99US-0145951.	PR	26-OCT-1999;	99US-0161360.
PR	02-AUG-1999;	99US-0146386.	PR	26-OCT-1999;	99US-0161361.
PR	02-AUG-1999;	99US-0146388.	PR	26-OCT-1999;	99US-0161920.
PR	02-AUG-1999;	99US-0146389.	PR	28-OCT-1999;	99US-0161992.
PR	03-AUG-1999;	99US-0147038.	PR	28-OCT-1999;	99US-0161993.
PR	04-AUG-1999;	99US-0147204.	PR	29-OCT-1999;	99US-0162142.
PR	04-AUG-1999;	99US-0147302.	Query Match 8.5%; Score 136.6; DB 21; Length 462;		
PR	05-AUG-1999;	99US-0147192.	Best Local Similarity 62.5%; Pred. No. 4.6e-24;		
PR	05-AUG-1999;	99US-0147260.	Matches 283; Conservative 0; Mismatches 134; Indels 36; Gaps 3;		
PR	06-AUG-1999;	99US-0147303.			
PR	06-AUG-1999;	99US-0147416.			
PR	09-AUG-1999;	99US-0147493.	QY	100	AAAGTCAAAGTCTCTGAAAAAACAACCGTGAACACACATAAACCC----ACTAGGAAAAAAA 156
PR	09-AUG-1999;	99US-0147935.	Db	4	agataaagattatgagcaaacacatgtgaaccacaaacacagctactactagggaagaaa 63
PR	10-AUG-1999;	99US-0148171.	QY	157	GAGTGTCAATTGGTAAACATTGGATCTTCTACCTAGCTTCTTATTACACCAAAAATTT 216
PR	11-AUG-1999;	99US-0148319.	Db	64	caattccacctcaccacatttgatcttccttctcctcactcacaacagagttt 123
PR	12-AUG-1999;	99US-0148341.	QY	217	CTCATCTATAAAA-----GGTGTGAAAACTTTGACGAGACG 252
PR	13-AUG-1999;	99US-0148565.	Db	124	ctactctacaagttccagaaccttcttagatctcgaggaaccacacattccaaaacagggtc 103
PR	16-AUG-1999;	99US-0149368.	QY	253	GTGGAAAAAATTAAAGATGGACTGGCGCTTAGTATTGGTGGATTCTTATCAACTAGCTGGG 312
PR	17-AUG-1999;	99US-0149175.	Db	184	gtggagaatctcaagatggttgggttcttggaagactctctatcaactcgcggt 243
PR	18-AUG-1999;	99US-0149426.	QY	313	AAACTTGGAAAAAGATCAAGAAGGGGTTTTTCAGGTGGAATACCGCAT-----GAC 363
PR	20-AUG-1999;	99US-0149723.	Db	244	aegctggccaagacgatgaaggagctctccgtgttgtagtaacgacgctgaagactcagag 303
PR	20-AUG-1999;	99US-0149902.	QY	364	ATGGATGGTGTAGAGGTGACAGTGGCTGTTGCAGAGAGATAGAAGTTGCAGATCTTACT 423
PR	23-AUG-1999;	99US-0149902.			
PR	23-AUG-1999;	99US-0149930.			
PR	25-AUG-1999;	99US-0150566.			
PR	26-AUG-1999;	99US-0150884.			
PR	27-AUG-1999;	99US-0151065.			
PR	27-AUG-1999;	99US-0151066.			
PR	27-AUG-1999;	99US-0151080.			
PR	30-AUG-1999;	99US-0151303.			

PR	09-AUG-1999;	99US-0147493.	QY	359	ATGACATGGTGGTGTAGAGGTGACAGTGGCTGTTGCAGAAGAGATAGAAGTTGCAGATC	418
PR	09-AUG-1999;	99US-0147935.				
PR	10-AUG-1999;	99US-0148171.	Db	310	acgatagcctcgcggggtttattctacgcgcacctcagatttatgtgtagaagatattcg	369
PR	11-AUG-1999;	99US-0148319.				
PR	12-AUG-1999;	99US-0148341.	QY	419	TTACTGATGAAGAAGCACCACCAATTCAGAGCATTTGATTCCTTCTAATAAATCTTGA	478
PR	13-AUG-1999;	99US-0148565.				
PR	13-AUG-1999;	99US-0148684.	Db	370	tgggtgctaagtatgcccccttcaattgttcaatcttcttaccaccacaaagcgtga	429
PR	16-AUG-1999;	99US-0149368.				
PR	17-AUG-1999;	99US-0149175.	QY	479	ATTTGGAAGGGCTTCATCGCCCTCTTCTAGCTGTGAGCTCACCACAGCTCAAGGACGGGC	538
PR	18-AUG-1999;	99US-0149426.				
PR	20-AUG-1999;	99US-0149722.	Db	430	atcatgatggacacacccatgagcctcttaccggtccaggtgacagatctggtagatggaa	489
PR	20-AUG-1999;	99US-0149723.				
PR	20-AUG-1999;	99US-0149929.	QY	539	TCACCATGGATTAGCATTTAACCATGCTGTGTGTGATGCTACTTCGACGTGCGACTTTA	598
PR	23-AUG-1999;	99US-0149902.				
PR	23-AUG-1999;	99US-0149930.	Db	490	tittcataggactgtctatgaaccatgcgatgggagacggttactgcgttcttgggaagtctt	549
PR	25-AUG-1999;	99US-0150566.				
PR	26-AUG-1999;	99US-0150884.	QY	599	TGACCTCGTGGTCCGAGCTTTCGCTGGGGTCCACCTCAATTTCTGCCACCATTCCTTG	658
PR	27-AUG-1999;	99US-0151065.				
PR	27-AUG-1999;	99US-0151066.	Db	550	ttaccgcatggctcgaatcttccaaggaacagagagcaaaaaacaatgacttgtgtc	609
PR	27-AUG-1999;	99US-0151080.				
PR	30-AUG-1999;	99US-0151303.	QY	659	AACGAACCAAGGCTCGTAACACTCGAGTCAAGCTCAACCTCTCTCAACCATCAGATGCCAC	718
PR	31-AUG-1999;	99US-0151438.				
PR	01-SEP-1999;	99US-0151930.	Db	610	ttaagaatccaccgcgtcttazagcgttacatccctgaagatacgcgtctctattcagcc	669
PR	07-SEP-1999;	99US-0152363.	QY	719	CGGAACATGCTAAGTCAGCAACCAACGGTGTATGCCGGCCAACTAGACCCACCTCTTC	778
PR	13-SEP-1999;	99US-0153758.				
PR	15-SEP-1999;	99US-0154018.	Db	670	ttccatatagccatccag-----atgaatttatcagaacacctacgaatctccaatcctca	723
PR	16-SEP-1999;	99US-0154039.				
PR	20-SEP-1999;	99US-0154779.	QY	779	GCGAAAGAGTATTCAAGTTCCTCGAGTTAGCAATTCACAAATTCAGTCAACAGTCAATG	838
PR	22-SEP-1999;	99US-0155139.				
PR	23-SEP-1999;	99US-0155486.	Db	724	aggagagaatgtctgtttctctatcagaaccataagattgcttaaaacaaagggtcaacc	783
PR	24-SEP-1999;	99US-0155659.				
PR	28-SEP-1999;	99US-0156458.	QY	839	CCNACTCAGGAGAGAGCGCCATTCCTCCACATTCCAATTCACCTCCGCGACACGTTGCTAG	898
PR	29-SEP-1999;	99US-0156596.				
PR	04-OCT-1999;	99US-0157117.	Db	784	aaatatgtggaacaaacctccatctcatcttccaatcgttaaccgcggttatatggagat	843
PR	05-OCT-1999;	99US-0157753.				
PR	06-OCT-1999;	99US-0157865.	QY	899	CGGTACACGTGCGCGCCCAACTCAAGCCCGAGGACTACACTGTGTACACTGTGTGCTG	958
PR	07-OCT-1999;	99US-0158029.				
PR	08-OCT-1999;	99US-0158232.	Db	844	gcataacaaggcgagagagattacacctcgcatacgcagagaacaagtgttagagttgcgtcg	903
PR	12-OCT-1999;	99US-0158369.				
PR	13-OCT-1999;	99US-0158293.	QY	959	ATTGCGAGAAAGGGTTGATCTCCAAATGCCAGAAATTCATTCGCGCAACCTTAATTCAGG	1018
PR	13-OCT-1999;	99US-0159294.				
PR	13-OCT-1999;	99US-0159295.	Db	904	ataacagaggaagaaatgtatccaccgcttcacaaggattacttcggaacctgcctctcg	963
PR	14-OCT-1999;	99US-0159329.				
PR	14-OCT-1999;	99US-0159330.	QY	1019	CAATTTTCACGTGACCGCGGAGGTTTGTACTACTACCAAGCCCGATCGAGTTCGCTGGTG	1078
PR	14-OCT-1999;	99US-0159331.				
PR	14-OCT-1999;	99US-0159637.	Db	964	cttggagaacgctgcataaaagctggtgaactattggagaatgatctctggagctgctt	1023
PR	14-OCT-1999;	99US-0159638.				
PR	18-OCT-1999;	99US-0159584.	QY	1079	GGATGATCAACAACGCGATCTGTAAGCATGACGCTAAGGCCATTGATGAAGAACAAGG	1138
PR	21-OCT-1999;	99US-0160741.				
PR	21-OCT-1999;	99US-0160767.	Db	1024	tgaagtgcatacaagcagtaigtgaacatacagagcgaaggatctcacaatgatgatc	1083
PR	21-OCT-1999;	99US-0160768.				
PR	21-OCT-1999;	99US-0160770.	QY	1139	AGTGGAGAGCAACCCGGAAGTCTTTTCAGTACAAAGAT---GCTGGAGTGAATGTGTTG	1195
PR	21-OCT-1999;	99US-0160814.				
PR	21-OCT-1999;	99US-0160815.	Db	1084	aatggctaaagtctctctatttaccattatgatcgcttatcttgaaccgatgagcgta	1143
PR	22-OCT-1999;	99US-0160980.				
PR	22-OCT-1999;	99US-0160981.	QY	1196	CTGTTGGAAGTTCGCCAAGGTTCAAGGTTTACCACGTGGATTTGGATGGGGAAGCCAG	1255
PR	22-OCT-1999;	99US-0160989.				
PR	25-OCT-1999;	99US-0161404.	Db	1144	tgatgggaagctgcgcgaggttataaagtatggttgaatttggagtggaaggag	1203
PR	25-OCT-1999;	99US-0161405.				
PR	25-OCT-1999;	99US-0161406.	QY	1256	AGAGTGTGAGGAGTGGTTCGMAACAATAGGTTTGTATGGTATTTGTACCAAGGCA	1315
PR	26-OCT-1999;	99US-0161359.				
PR	26-OCT-1999;	99US-0161360.	Db	1204	tgaagcttagaagtggtctatgcataaagtctgatggaaggatctcttaccacgggaa	1263
PR	26-OCT-1999;	99US-0161361.				
PR	28-OCT-1999;	99US-0161920.	QY	1316	AAAAATGGAGGAAGAAGCATTTGATGTGGAGATTAGTTTGGAAAGCAAAATGCTAGGAGGT	1375
PR	28-OCT-1999;	99US-0161992.				
PR	28-OCT-1999;	99US-0161993.	Db	1264	gagaaggagggaagataaagtattggaggatgtctctgttccagagatttatggaagcctt	1323
PR	29-OCT-1999;	99US-0162142.				
PR	29-OCT-1999;	99US-0162142.	QY	1376	TGGAGAAAGATAAAGAGTTCTT	1397
PR			Db	1324	tggaaatcadatgaagagttcat	1345

Query Match 8.2%; Score 132; DB 21; Length 1363;
Best Local Similarity 46.8%; Pred. No. 9.6e-23;
Matches 488; Conservative 0; Mismatches 545; Indels 9; Gaps 2;

RESULT 10
AAC42254
ID AAC42254 standard; DNA; 1520 BP.
XX
AC AAC42254;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 34861.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
FD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 18-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 11-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.

[illegible]

RESULT	14
AAF58259	
ID	AAF58259 standard; DNA; 936 BP.
XX	
AC	AAF58259;
XX	

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:38:25 ; Search time 8904.87 Seconds
(without alignments)
2787.883 Million cell updates/sec

Title: US-08-894-356C-3
Perfect score: 1605
Sequence: 1 TGTGACGAAATCCATTTC.....TTAAAAAAAAAAAAAAAAAAAA 1605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl : *
1: gb_ba1 : *
2: gb_ba2 : *
3: gb_ba3 : *
4: gb_in1 : *
5: gb_in2 : *
6: gb_in3 : *
7: gb_om : *
8: gb_ov : *
9: gb_pat1 : *
10: gb_pat2 : *
11: gb_ph : *
12: gb_pl1 : *
13: gb_pl2 : *
14: gb_pl3 : *
15: gb_pl4 : *
16: em_ba1 : *
17: em_ba2 : *
18: em_fun : *
19: em_htgo_hum : *
20: em_htgo_inv : *
21: em_htgo_rod : *
22: em_htg_hum1 : *
23: em_htg_hum2 : *
24: em_htg_hum3 : *
25: em_htg_hum4 : *
26: em_htg_hum5 : *
27: em_htg_hum6 : *
28: em_htg_hum7 : *
29: em_htg_hum8 : *
30: em_htg_inv1 : *
31: em_htg_inv2 : *
32: em_htg_other : *
33: em_htg_rod : *
34: em_hum1 : *
35: em_hum2 : *
36: em_hum3 : *
37: em_hum4 : *
38: em_hum5 : *
39: em_hum6 : *
40: em_hum7 : *
41: em_in : *
42: em_om : *
43: em_or : *

44: em_ov : *
45: em_pat : *
46: em_ph : *
47: em_pl : *
48: em_ro : *
49: em_rod : *
50: em_sy : *
51: em_un : *
52: em_vi : *
53: gb_sts1 : *
54: gb_sts2 : *
55: gb_sts3 : *
56: gb_sy : *
57: gb_un : *
58: gb_vil : *
59: gb_vil2 : *
60: gb_vil3 : *
61: gb_vil4 : *
62: gb_vil5 : *
63: gb_vil6 : *
64: gb_vil7 : *
65: gb_vil8 : *
66: gb_vil9 : *
67: gb_vil10 : *
68: gb_vil11 : *
69: gb_vil12 : *
70: gb_vil13 : *
71: gb_vil14 : *
72: gb_vil15 : *
73: gb_vil16 : *
74: gb_vil17 : *
75: gb_vil18 : *
76: gb_vil19 : *
77: gb_vil20 : *
78: gb_vil21 : *
79: gb_vil22 : *
80: gb_vil23 : *
81: gb_vil24 : *
82: gb_vil25 : *
83: gb_vil26 : *
84: gb_vil27 : *
85: gb_vil28 : *
86: gb_vil29 : *
87: gb_vil30 : *
88: gb_vil31 : *
89: gb_vil32 : *
90: gb_vil33 : *
91: gb_vil34 : *
92: gb_vil35 : *
93: gb_vil36 : *
94: gb_vil37 : *
95: gb_vil38 : *
96: gb_vil39 : *
97: gb_vil40 : *
98: em_ba3 : *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1603.4	99.9	1605	10	E12755	E12755 Petunia hyb
2	1603.4	99.9	1605	12	AB026495	AB026495 Petunia x
3	378.8	23.6	82415	12	AB005244	AB005244 Arabidops
4	135.2	8.4	79976	12	AB008264	AB008264 Arabidops
5	131.8	8.2	118507	13	ATF7J8	AL137189 Arabidops
6	119.2	7.4	83599	12	AB005249	AB005249 Arabidops
7	119.2	7.4	88095	13	ATF13G24	AL133421 Arabidops
8	106.2	6.6	110531	12	AF002109	AF002109 Arabidops


```
QY 541 ACCATGGATTAGCATTTAACCATGCTGCTGCTGATGCTACTTCGAGGTGGCACTTTATG 600
|
|
|
Db 541 ACCATGGATTAGCATTTAACCATGCTGCTGCTGATGCTACTTCGAGGTGGCACTTTATG 600
|
|
|
QY 501 ACCTCGTGGTCCGAGCTTGTGCTGGGTGCACCTCAATTTCTGTCCACCACTTCCTTGAA 660
|
|
|
Db 501 ACCTCGTGGTCCGAGCTTGTGCTGGGTGCACCTCAATTTCTGTCCACCACTTCCTTGAA 660
|
|
|
QY 561 CGAACCAAGGCTCTAACAATCGAGTCAAGCTCAACCTCTCTCAACCATCAGATGCACCC 720
|
|
|
Db 561 CGAACCAAGGCTCTAACAATCGAGTCAAGCTCAACCTCTCTCAACCATCAGATGCACCC 720
|
|
|
QY 721 GAACATGCTAAGTCAAGCAACCAACCGGTGATGCTCCCGGCCAACGCTAGACCACTCTTCGC 780
|
|
|
Db 721 GAACATGCTAAGTCAAGCAACCAACCGGTGATGCTCCCGGCCAACGCTAGACCACTCTTCGC 780
|
|
|
QY 781 GAAAGAGTATTCAAGTTCTCGGAGTTAGCAATTTGACAAAATCAAGTCAACAGTGC 840
|
|
|
Db 781 GAAAGAGTATTCAAGTTCTCGGAGTTAGCAATTTGACAAAATCAAGTCAACAGTGC 840
|
|
|
QY 841 AACTCAGAGAGAGCCATTCTCCACATTCCAATCACTCTCCGCACACGCTGGCTAGCC 900
|
|
|
Db 841 AACTCAGAGAGAGCCATTCTCCACATTCCAATCACTCTCCGCACACGCTGGCTAGCC 900
|
|
|
QY 901 GTCACAGTGGCGGCCCACTCAAGCCGAGGACTACACTGTGTACACTGTGTTTCTGCTGAT 960
|
|
|
Db 901 GTCACAGTGGCGGCCCACTCAAGCCGAGGACTACACTGTGTACACTGTGTTTCTGCTGAT 960
|
|
|
QY 961 TGCAGGAAAGGGTTGATCTCCAAATGCCAGAAAGTTACTTCGGCAACCTTAATTCAGGCA 1020
|
|
|
Db 961 TGCAGGAAAGGGTTGATCTCCAAATGCCAGAAAGTTACTTCGGCAACCTTAATTCAGGCA 1020
|
|
|
QY 1021 ATTTTCACAGTACCGCGGAGGTTTCTTACTAGCAAGCCGCTGAGTTCGCTGGTGGG 1080
|
|
|
Db 1021 ATTTTCACAGTACCGCGGAGGTTTCTTACTAGCAAGCCGCTGAGTTCGCTGGTGGG 1080
|
|
|
QY 1081 ATGATACAACAAGCGATCGTGAAGCATGACGCTAAGGCCATTTGATGAAGAAACAAGGAG 1140
|
|
|
Db 1081 ATGATACAACAAGCGATCGTGAAGCATGACGCTAAGGCCATTTGATGAAGAAACAAGGAG 1140
|
|
|
QY 1141 TGGAGAGAACCCGAGATCTTTCAGTACAAGATGCTGAGTGAAGTGTGTTGCTGTT 1200
|
|
|
Db 1141 TGGAGAGAACCCGAGATCTTTCAGTACAAGATGCTGAGTGAAGTGTGTTGCTGTT 1200
|
|
|
QY 1201 GGAAGTTTCGCAAGGTTTCAAGTTTACGAGTGGATTTTGGATGGGAAAGCCAGAGAGT 1260
|
|
|
Db 1201 GGAAGTTTCGCAAGGTTTCAAGTTTACGAGTGGATTTTGGATGGGAAAGCCAGAGAGT 1260
|
|
|
QY 1261 GTGAGGAGTGGTTCGAACAATAGGTTTGTGATGGAATGGTGTATTTGTACCAAGGCAAAAAT 1320
|
|
|
Db 1261 GTGAGGAGTGGTTCGAACAATAGGTTTGTGATGGAATGGTGTATTTGTACCAAGGCAAAAAT 1320
|
|
|
QY 1321 GGAGAGAGAGCATTTGATGGAGATAGTTTGAAGCAAAATGCTATGGAGAGTTGAG 1380
|
|
|
Db 1321 GGAGAGAGAGCATTTGATGGAGATAGTTTGAAGCAAAATGCTATGGAGAGTTGAG 1380
|
|
|
QY 1381 AAGATAAAGAGTTCCTTCATGGAACCTGCTTAATTTGCTTACGTTGAGCTCAACTGGCTA 1440
|
|
|
Db 1381 AAGATAAAGAGTTCCTTCATGGAACCTGCTTAATTTGCTTACGTTGAGCTCAACTGGCTA 1440
|
|
|
QY 1441 CACTTTTATTTAGAGTGCCTATGACTCACAATGCATGATGTTATTTTGGAGGGGT 1500
|
|
|
Db 1441 CACTTTTATTTAGAGTGCCTATGACTCACAATGCATGATGTTATTTTGGAGGGGT 1500
|
|
|
QY 1501 TCCTTCCTTTTATGTTTCTATGTTTCTTCTTCTGTCAGTTATCAAGAGAAACCGAG 1560
|
|
|
Db 1501 TCCTTCCTTTTATGTTTCTATGTTTCTTCTTCTGTCAGTTATCAAGAGAAACCGAG 1560
|
|
|
QY 1561 TATAAGGAATAATGTTTTCAGTTTATTAAGGAAAAAAGAAAAA 1605
|
|
|
Db 1561 TATAAGGAATAATGTTTTCAGTTTATTAAGGAAAAAAGAAAAA 1605
|
|
|
```

```
RESULT 3
AB005244/c
LOCUS
DEFINITION
AB005244 BAO00015
VERSION
AB005244.2 GI:10176838
KEYWORDS
SOURCE
Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
clone:MRO11
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.
1 (sites)
Sato,S., Kotani,H., Nakamura,Y., Kaneko,T., Asamizu,E., Fukami,M.,
Miyajima,N. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 5. I.
Sequence features of the 1.6 Mb regions covered by twenty
physically assigned P1 clones
DNA Res. 4 (3), 215-230 (1997)
97471969
2 (bases 1 to 82415)
Nakamura,Y.
Direct Submission
Submitted (02-JUL-1997) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
On Sep 15, 2000 this sequence version replaced gi:2264316.
Address for correspondence: kaos@kazusa.or.jp
http://www.kazusa.or.jp/kaos/cgi-bin/agd/graph.cgi?c=MRO11
For the latest information on annotation of this clone, please see
http://gremlin1.zoel.iastate.edu/cgi-bin/sp.cgi).
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremlin1.zoel.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is MQM1 and the 3' clone is MZF18.
FEATURES
Location/Qualifiers
1..82415
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone_lib="Mitsui P1"
/clone="MRO11"
2972..4078
/clone_lib="Mitsui P1"
2972..4078
/note="contains similarity to photomorphogenesis repressor
protein
gene_id:MRO11.23"
/codon_start=1
/evidence-not_experimental
/protein_id="BABI0046.1"
/db_xref="GI:10176840"
CDS
```


Qy	694	AACCTCTCTCAACCATCAGATGCACCGAACAATCTAAGTCAGCAACCAACGGTGATGTC	753
Ds	75123	GACCTACTGCTCCCAAGGACCCCAACGAGACCTCCACGCGGAA---GATGCTGCCAAT	75067
Qy	754	CCGGCCAAAGTAGACCAACCTCTTCGGGAAAGAGATATCAAGTTCTCCGAGTTAGCAATT	813
Ds	75066	CCGACGGTGGAAACCGCGGACCTAGTAGAAGAATCTTACGTTCTCGGACTTTCGCTGTC	75007
Qy	814	GACAAATCAAGTCAACAGTCAATGCCAATCTCAGAGAG-----ACGCCAATTCTCC	864
Ds	75006	CACAGTCAAGTCAAGAGCTAAATTCAGTAAATCCATCGGACAGTTCAAAACCATCTTCA	74947
Qy	865	ACATTCCAATCACTCTCCGACAGCTGTGGCTAGCCGTACACAGTGGCGGCAACTCAAG	924
Ds	74946	ACATTCCAATCCCTGACATCACATCATCTGGCGCCACGTCACCTTAGCGGTGGACTCAA	74887
Qy	925	CCGAGGACTACACTGTGACACTGTGTTGCTGATTCAGAGAAAGGTTGATCCTCCA	984
Ds	74886	CCGAGGAGACATAAATATTTTCCACGCTCTCCGCGACTCTCCGCCCGCGTTGATCCTCCA	74827
Qy	985	ATGCCAGAAAGTTACTTCCGCAACCTAATTCAGGCAATTTTTCAGAGTACCGCGCGAGGT	1044
Ds	74826	ATGCCGAGGAATATTTCCGAAACTTGATTCACGAGATCTTTCACCGGAACAGCGCGGG	74767
Qy	1045	TTGTTACTAGCAAGCCGATCGAGTTCGCTGGTGGGATGATACAAACGATCGTGAAG	1104
Ds	74766	CTACTAGCAGCGCAGGACCGAGTTCGGAGCTTCGGTGATCCAGAAACATCGCGCA	74707
Qy	1105	CATGACCGTAAAGCCATGTATGAAGAAACAAAGAGTGGGAGACACCCGAAGATCTTT	1164
Ds	74706	CACGACCGAGTGTGTCGACGACGCTAACGACGAGTGGGAGAAATGCCCGAAGATTTT	74647
Qy	1165	CAGTACAAAGATGCTGGAGTCAACTGTCTGCTGTCGAGTTCGCAAGGTTCAAGTT	1224
Ds	74646	CAATTCAAAGACCGGAGTGAATTTGCTCGGTGGGAAGCTCACCGAGTTTCGTT	74587
Qy	1225	TACGACCTGGATTTTGGATGGGAAAGCAGAGAGTGTGAGAGTGTGCAACAATAGG	1284
Ds	74586	TACGAGGTGGATTTTCGGGTTTGGAAACCGGAACGCTTAGGACGGTCAATTAACCG	74527
Qy	1285	TTTGAATGAATGTTGATTTTACACAGGCAAAATGAGGAAGAACGATGATGTGGAG	1344
Ds	74526	TTTAAATGGATGATGATTTTCTACCAAGGTAAGGTAAGCAGGATTAAGTATTGATGTGG	74467
Qy	1345	ATTAGTTTGGAAAGAAATGCTATGAGAGGTTGGAGAAAGATAAAGAGTT	1394
Ds	74466	ATACTCTGAAGCTTCTGTTATGAGAGAGCTTGTGAAGACCAAGAGTT	74417
RESULT	4		
LOCUS	AB008264		
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, p1 clone:MBD2.		
ACCESSION	AB008264	BA000015	
VERSION	AB008264.1	GI:2618599	
KEYWORDS	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui p1 clone:MBD2.		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (sites)		
AUTHORS	Nakamura,Y., Sato,S., Kaneko,T., Kotani,H., Asamizu,E., Miyajima,N. and Tabata,S.		
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. III. Sequence features of the regions of 1,191,918 bp covered by seventeen physically assigned p1 clones		
JOURNAL	DNA Res. 4 (6), 401-414 (1997)		
MEDLINE	96162728		
REFERENCE	2 (bases 1 to 79976)		
AUTHORS	Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (24-OCT-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)		
COMMENT	Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/seqd.graph.cgi?c=MBD2 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlin.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MJB21 and the 3' clone is MRD20.		
FEATURES	Location/Qualifiers		
SOURCE	1. 79976 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /chromosome="5" /clone="MBD2" /clone_lib="Mitsui p1" complement(1..218) /note="CDS is reported in Acc# AB007647 gene_id:MJB21.20" /number=1 /evidence=not_experimental /product="U2 snRNP auxiliary factor, small subunit" join(5306..5761,5932..6828) /note="gene_id:MBD2.2" /codon_start=1 /evidence=not_experimental /product="N-hydroxycinnamoyl/benzoyltransferase-like protein" /protein_id="BAB09184.1" /db_xref="GI:9758571" /translation="MDSSSVKIVSKSVKPKTLPEESKQPYLSPWDYAMLVSQYIQ KGLLFHKPLDSDITLLEKLDLAVTLVHFYPLAGRLSLATTEKPKRSVFFVDCNDS PGAGFIATSDLCIKDIVGAKYVPSIVQSFDDHKA VNHGHTMSLSISVQVTELVGDI FIGLSNNHAMDGTAFWKPTAFWSEIFQGESQNDLCLKNPVLKRYIPEGYGLF SLPSHPDEPIRTVESPILKRMFCFSETIRMLKTRVNOICCTTSSPQSITAVTW RCITRARRPLDRETSCTRNAADNRMYPLHKDYFGNCLSLARTAKAGELLENDLG FAALKVQVAEHTSEKVSQMDWLKSPYIHLRLEFPMVSMGSSSPFNKYGCEF GLGKGVTLRSYAHKFDGVSAYPGREGGSDIDLECLVPEFMEALDESDEERFLSVSL "		
CDS	8080..10095 /note="gene_id:MBD2.3" /codon_start=1 /evidence=not_experimental /product="CHP-rich zinc finger protein-like" /protein_id="BAB09185.1" /db_xref="GI:9758572" /translation="MSLYFEGHEHHSIIKRDGLSCDADRSFGDVISGCEKFTVH RKCFMFQDEIFDHPSHDGHCLLTITGAPDITDQCHLCGRKTKRLLYHSDCKLN VTDICIDYICARSPKLPMDHDLIKVDLGANNPCDFCNESGIDPCRCRPMIHER CASVDFSPETTHPSHVHPLKLSNGAPDTNLNCHTCGDTANGLYHCDICRFLNRM CAIRAEPTVALSNMKVHEHTLMPRLISFVCDACMGKDRAPYVCQCDNFIHOE CAKLPRIVHNVHHRVSYKPLGPGECWGCWEEDWSYGAYSCSLCPNVAHSLC ATRKDVNDGRELGVLEEDIEDIEPFKRNDDNTITHTHEHNLKSDGITLAKSILVAC VCPIDGTFYNSSESSCFVLSEKSRKHFLLSPVLCLQNQRNVTFCNACQ QVFEKGFIVSSFPKTVYRKGFYDLQCSSTVFFHSGHSHLLFLKLGKRNVTCCBC GIVQKGYAIGCTKCNVFLDFRCATLPLTVRLPRYDDHPLTLCYGDEKASGKCWDCICE		

CDS
RETNPKSWFYTKDCGVTFHIFCVVMDIRFSKQGETIRDGVELLPNNSTSRPLCKNCQ
CRCLGPFVKQDNYCYSYCYARLHSLRWISKRCPPWISSEPT*
complement(join(10994..11250,11343..11490))
/note="gene_id:MBD2.4
unknown protein"

/codon_start=1
/evidence=not_experimental
/protein_id="BAB09186.1"
/db_xref="GI:9758573"
/translation="MTLKKVDANSTLESSLOELKSDTSSRSKINFILFLADNDPTTG
QSMPCDVRABEPIVKTLEEPVEVKLIRAYAGDRPTWRTFAHPMRVDSRFLKLGVPT
LVRWGDGVKGLRLEDHQHLPLHLPLAPST"
complement(join(12261..12611,13086..13271,13354..13779))
/note="emb|CAB53482.1
gene_id:MBD2.5
similar to unknown protein"

/codon_start=1
/evidence=not_experimental
/protein_id="BAB09187.1"
/db_xref="GI:9758574"
/translation="MHAKTDSVTSLSASSPTRPRPAYFVQSPSRSDSHDGEKTATS
FHSPTVLTSPGSPHSHSSSRFSKINGSRRKHAGCKQFAMIEEQLLDGDRQE
ALPRCYVIAIVGLFLLFAFSLILYAAAPQPKISVKRSITFEQLKVQAGQDAGGI
GTMIDMTNATLMRLYRNITFFGVHVTSPIDLFSQITIGSGIKKPKYQKRSQRTV
VYVNLGDKPIYLGSGSTLVPPPPAPIPKPKKKGPVIVPPPPAPVPMRLNFTVR
SRAYVLGKLVOPKFKYKRVCLINFEHKKLSKHPITNNCTVSI"
complement(join(14641..14694,14770..14868,15144..15272,
15358..15401,15497..15569,15650..15736,15848..15932,
16050..16123,16204..16733,16863..17553,17677..18603))
/note="gb|AAFP23287.1
gene_id:MBD2.6
strong similarity to unknown protein"

/codon_start=1
/evidence=not_experimental
/protein_id="BAB09188.1"
/db_xref="GI:9758575"
/translation="MNAYGRIGCSYIYRGVTVSGPFPFGGAIDIIIVVEQDGTGFKSS
PWYVRFGQVLTKRNLRIDVNGVDSGFNMYLAHTGOAVFLREVEDVVGESGSE
VYTLSSGAEVLTSRDQVVKVPLKSRNCNDVSPSPRTNGKTVLGVKPLIGLVYFGG
RSVRESQDCGVAREIAADLLEVKNSTNIDTRKRGKMSSESLDKDKGESTSTSGKS
CVCESAMVVSQDSILEPLVASPTLRLDEKEQDFRESTINVEDYCEBDSGVVVEN
GLGSAASVFTSVSEGSNVEIFVEPRTEALAEADVAGSDLOSLOELLRAPESVIAI
LGSADQMSGITGTSQEGSGSGSPVDENKITIKDMHISAGDFEKSQASGESILQE
IEE0FSPSDDLECKPGNSGSSSDTVKVDGESYDETKTSPKGVENTMALSEP
INIERKDDIFDMEERLVCSLIMRLONNDMDASPSQPLSQSEDPDPCNTSKLDRLE
ESSGGLDAESVENSPLKAFKHVIANPEVVELSLCKHLISEGNGACAAQAFNSEK
LDMEKFSGLSEILNKLVKIGCYEPMDAAAPLILGVSVFQVFEKPKVIAVD
RNEKPGVLAQSGSKWLMWFLSLRRSTKEAESPSGDTPAEPEEKOEKSPRPMKTVR
ALPTTSEQLASLDLDGMNSVTFTFTNIVGTQQVDARIYLWKNRSRVIVSDVGTIT
RSDVLGQPMPLVGDMSOTGVTHLFSVAVKNGYQLIFLSARAIISOAVTROPFLVNLQ
DGKALPDGPVVISPDLPSPSLFREVIRRAPHEFKIACLEEIRGLPPEHNPYAGFGN
RDTDEISLVKVGIPRGKIFINPRGEVAVNRIDTRSYTNLHTLVNRFPATSSPEPE
DFTWNWNLKLPPLSLM"
Join(20680..21174,21253..23013)
/note="gene_id:MBD2.7
pir||T00990
strong similarity to unknown protein"

/codon_start=1
/evidence=not_experimental
/protein_id="BAB09189.1"
/db_xref="GI:9758576"
/translation="MLGDDKDDDLNLFUNAIGEAGDEGPTSFNDIDFLPDDRDLH
NPQDCEPTNNEPKVYVAPRVMISHODSFSEDSRTDVIEDARTLPASRLRVPASP
RAFYVPRSVESPRFGSPRSVSPCGSPIGVIDTASPVESVREAVSKFGGIDTWKAKH
IQTIBRRKWDQDELEKIQEAMPEYKRAELAEAAKIDALEELENTKGLIEELKLELEK
AEKEQAKQSELAQVMEKEGVANESAVVKTLEAVAKARQVSTSELRSVREE
ITEMVNSYKMLRADIAERADIAVLEAKETERTMDGLSIELIAKELLESVHPAHL
EAEKRFVSMARQDQVYVWEKELMKNVDIERNQEVRAADDVKAKLVEFSALQHD
KTELAFDIDISSGNLLLEKNDIHAAVESARRELEVKANIEKAASEVKKLIIAGSLQ
SELGRERDLETKQESTGLARTNDKQGBELVETAKKLQATKEADAKATASR
DELWNAKLSQAQKMGSTIESRLVYAKEMEAARSAKLLAALAIKALQGETSSORFE
EINNSPRITISVEYVYELSQALESSEANTRLSEIVSQIEVAKEEESRILEKEEV
NREMSYKAELEKANGKAEKARDKGLGMEQELKWRSENGKRRTDEGREPEKSPTRSS

CDS
TEGRNKGFGQSKSFARFGEQSSSNNTGGSTTTNNNLTPETKKKKKKLSLFPKVF
FLSRKSHSHK*
complement(join(23380..23461,24198..24300,24565..24625,
24707..24832))
/note="gene_id:MBD2.8
pir||T28058
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09190.1"
/db_xref="GI:9758577"
/translation="MANTQLKSDAIMDMMKHLSLDAGKEVTEKILGVYQINIAPKKL
GFEEVTVLDLKGVEVTKYEGKVDFATFSKDDDFVKVATGKMNPMQAFIRGAMKI
KGSLSAAOKETPDIFPKPSKL"
complement(join(25760..25834,27425..27528,27615..27900,
27989..28317,28527..28634))
/note="gene_id:MBD2.9
pir||T05226
similar to unknown protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB09191.1"
/db_xref="GI:9758578"
/translation="MVGDIYRGFRSRRFSDSDSDSDASSVEGETTSSMSYAGKEYM
ETETWNEKHSLYLKSMEASFVDQLYNSLGAICKNENVESTFRGSGRKPQSQKVLH
DGFQKINVQKQPEHRINGRNGSHDFLRSWIKHYKPLVKTOIVTDEPENOVVSS
NGKKGICSSGASSLKQLSSHSRDHDIQSVGEEVSDQNFVNEGKINGESSKKMKTV
MSESSDQESVGTETEKETQDESVAIEHLQK*
complement(30319..31245)
/note="gene_id:MBD2.10"
/codon_start=1
/evidence=not_experimental
/product="non-LTR retroelement reverse transcriptase-like

CDS
Query Match 8.4%; Score 135.2; DB 12; Length 79976;
Best Local Similarity 48.5%; Pred. No. 2.1e-21;
Matches 439; Conservative 0; Mismatches 458; Indels 9; Gaps 2;
QY 495 TCGCCCTCTTAGCTGTCAGCTCACCAGCTCAAGGAGCGGCTCACCATGGGATTAGC 554
DB 5910 TCAGTCACATTTGTTGTAGGTGACAAATTTGGTAGATGGAATTTTCATAGGACTGTC 5969
QY 555 ATTTAACCATGCTGCTCGATGGTACTTTCGACGTGGCACCTTTATGACCTCGTGGTCCGA 614
DB 5970 TATGAACCATGCGATGGGAGACGGTACTCGGTCTCGAGTCTTTTACCGCATGGTCCGA 6029
QY 615 GCTTTGCTGTGGTCCACCTTAATTTCTGTCCACCATTTCTTGAACGAACCAAGGCTCG 674
DB 6030 AATCTTCCAAGGACAAGAGAGACAACCAAAACGATGACTTGTGTCTTTAAGAAATCCACCCGT 6089
QY 675 TAACACTCCAGTCAAGCTCAACCTCTCTCAACCATCAGATGCACCCGAACATGCTAAGTC 734
DB 6090 GTTGAAGCGTTACATCCCTGAGGATACCGTCTCTATTTCAGCCTTCCCATATAGTCATCC 6149
QY 735 AGCAACCAACGGTGTATGTCCTCCGCCAACCTTAGACCCACCTCTTCGCGAAGAGATTTCAA 794
DB 6150 AG-----ATGAATTTATCAGAACCTACGAATCTCCAATTTCTCAAGGAGAGATGTTCTG 6203
QY 795 GTTCTCCGAGTTAGCAATTTGACAAAATCAAGTCAACAGTCAATGCCCACTCAGGAGAGAC 854
DB 6204 TTTCTCATCAGAAACCATTAAGATGCTTAAACACAGGGCTCAACCAATATGCGGACAAC 6263
QY 855 GCATTCTCCACATTCCAATCACTCTCCGCACACAGTGTGCTAGCGTGCACACGTGCGG 914
DB 6264 CTCATCTCATCTTTCCAAATCGCTAACCCGGGTTATATGAGATGCTATACAAAGGCGGAG 6323
QY 915 CCAACTCAAGCCCGGAGCTACACTGTGTACACTGTGTTTGTGATTGAGGAAAAAGGCT 974
DB 6324 GAGATTACCTCTCGATCGAGAACAAAGTTGTAGAGTTGCTGCTGACACAGGAGAAGAT 6383
QY 975 TGATCCTCCCAATGCCAGAAGTTTACTTTCGGCAACCTTAATTCAGGCAATTTTTCACAGTAC 1034
DB 6384 GTATCCACCGCTTCAACAAGGATTACTTTCGGAACCTGCCTCTCTGCTTTGAGAACGCGTGC 6443

Qy	1035	CGCGCAGGTTTGTACTAGAACCCCGCATGAGTTCGCTGGTGGGATGATACAAACAGC	1094	gene	/number=1
Db	6444	AAAAGCTGGTGAACATATTGGAGAAATGATCTTGGATTGCTGCTTTGAAAGTGATCAAGC	6503	gene	/gene="F7J8_5"
Qy	1095	GATCGTGAAGCATGACGCTAGGCCATTGATGAAGAAACAGAGTGGGAGACCAACC	1154	CDS	complement(join(1726..2207,2629..2752,2857..2962,3091..3556,3633..3687))
Db	6504	AGTAGCTGAACATACGACGCAAAAGGTATCTCAATGATTGATCAATGGCTAAAGTCTCC	6563		/gene="F7J8_5"
Qy	1155	GAAGATCTTTCA--GTACAAAGATGCTGGAGTGAACCTGCTTCTGTTGGAAGTTCGCC	1211		/notes="strong similarity to serine/threonine-specific protein kinase NAK, Arabidopsis thaliana, EMBL:AL162874"
Db	6564	TTATATTTACCATATTGATCGGCTATTTGACCCGATGAGCGTTATGATGGGAAGCTCGCC	6623		/codon_start=1
Qy	1212	AAGGTTCAAGCTTTACGACGTTGGATTTTGGATGGGAAAGCCAGAGTGTGAGAGTGG	1271		/product="protein kinase-like protein"
Db	6624	GAGGTTTAATAAGTATGCTTGTGAATTTGGGTTCGGAAAGAGTGCACGCTTAGAAGTGG	6683		/protein_id="CAB99493.1"
Qy	1272	TTCGAACATAGGTTGATGAATGGTGTATTGTTACCAAGGCAAAATGAGGAGGAAG	1331		/db_xref="GI:9542860"
Db	6684	CTATCGCATAGTTCGATGAAAGGTATCTGCTTACCCGGGAAGAGAGGGGGAAG	6743		/translation="MGNCGTRDEAAVFTPOAAQOLQKKHSRSVSDLSDPSTPRFRDD
Qy	1332	CATTGATGTGAGATTAGTTTGGGAAGCAAAATGCTATGAGAGTTGGAGAAAGATAAGA	1391	intron	SRTPISYAQVIFETLFELETTTKSRPDDYILGEGFGTVYKGYIDNLRVLGSLPVA
Db	6744	TATAGATTGGAGGTAGTCTTGTTCAGAGCTTTATGGAAGCTTGGAAATCAGATGAAGA	6803	exon	KVKNREGGOGHREWLTEVNFGLQLRNPVLKLIYCCEDDHRLLLYEFMLRSLSNH
Qy	1392	GTTTCT 1397		intron	LFSNPRTLLHLKLMIALGAAGLAPLHNAERPVIYRDKFTSNLLSDSYTAKLSDF
Db	6804	GTTTCT 6809		exon	GLAKAPQGDGTHVSTRVMGTGYAAPEYMTGHLTARSVDYSGVVVLEMLTGKRSV
RESULT	5			intron	DKTRPSKEQNLVDWARPKLNDKRKLQIIDPRLNQYSVRAAGKAGSLAYICLSQNP
ATF7J8				exon	ARPLMSDVVFTLEPLOTCTGALLPCATTTTAGAFAAMGGVDPDYRMHRRFAKNVGP
LOCUS	ATF7J8	118507 bp	DNA	intron	RSPNPNYSGGPAACRV"
DEFINITION	Arabidopsis thaliana DNA chromosome 5, BAC clone F7J8 (ESSA			exon	/number=1
ACCESSION	AL137189			exon	complement(2629..2752)
VERSION	AL137189.2	GI:9542859		intron	/gene="F7J8_5"
KEYWORDS	thale cress.			exon	number=2
ORGANISM	Arabidopsis thaliana			exon	complement(2753..2856)
REFERENCE	1 (bases 3758 to 118507)			exon	complement(2857..2962)
AUTHORS	Wohlmann, P., Smith, A., Bancroft, I., Gruenisen, A., Wambutt, R., Kalicki, J., and Mayer, K.F.X.			intron	/gene="F7J8_5"
JOURNAL	Unpublished			exon	number=3
REFERENCE	2 (bases 1 to 3766)			exon	complement(2963..3090)
AUTHORS	Bevan, M., Rose, M., Hempel, S., Entian, K.-D., Bancroft, I., Mewes, H.W., Rudd, S., Lemcke, K. and Mayer, K.F.X.			exon	complement(3091..3556)
JOURNAL	Unpublished			intron	/gene="F7J8_5"
AUTHORS	3 (bases 1 to 118507)			exon	number=4
JOURNAL	EU Arabidopsis sequencing, project.			exon	complement(3557..3632)
AUTHORS	Submitted (26-JUL-2000) MIPS, at the Max-Planck-Institut fuer Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, FRG, E-mail: lemcke@mips.biochem.mpg.de, mayer@mips.biochem.mpg.de, Project Coordinator: Mike Bevan, Molecular Genetics Department, Cambridge Laboratory, John Innes Centre, Colney Lane, NR4 7UJ Norwich, UK, E-mail: michael.bevan@bbsrc.ac.uk			exon	number=5
TITLE	Direct Submission			exon	/gene="F7J8_10"
JOURNAL	On Jul 27, 2000 this sequence version replaced gi:6759425. Information on performance of analysis and a more detailed annotation of this entry and other sequences of chromosomes 3, 4 and 5 can be viewed at: http://www.mips.biochem.mpg.de/proj/thal/ .			exon	/gene="F7J8_10"
FEATURES	Location/Qualifiers			CDS	number=1
source	1..118507				join(6055..8082,8214..8420)
	/organism="Arabidopsis thaliana"				/gene="F7J8_10"
	/variety="Columbia"				/note="EST F14175 marks 5' end of this gene; 5' UTR contains an intron
	/db_xref="taxon:3702"				similarity to predicted proteins, Arabidopsis thaliana"
	/chromosome="5"				/codon_start=1
exon	complement(1726..2207)				/product="putative protein"
	/gene="F7J8_5"				/protein_id="CAB69831.1"
					/db_xref="GI:6759426"
					/translation="MAEKEKAGSGTNDNRVTKRSSHRRKCEVCNKSLEHNDLVKY
					MSKLPGLQRIERGESVHOSNVLVGVLDWESLQWKHGRAGKEISGRSERKVTI
					ATTSTGVVVPNDSANRCKTDDOVHTCSNLGVKASPDLOYLEPOLASRDLNKOET
					ATUCYSKSSGRDHKGVEPRKSRRTSHNRSTTGLSSEMGNSAGSLFRDKETQKAGEIH
					AKEARERAKCEVKLDGDEKLIIGDSEAGLTSEKQEFSENIPLLRKSKRSRSTLSEPOI
					SREVNRLSDFSDGINSFSLRSPSCPLSFLDERSDMDMLPLGTDLSLGRGGRKH
					SKTTSRTDFRFPEDSKRHPKSRFSFGRSLRSLRDLPLPKKASESVLPSPKARSSS
					MSGSMRFDGVCPSQSNPENQTHCRSVPLRRFLDPLKPKASESVLPSPKARSSS
					SNPKPTNSNVLQDEKKQDASRTLAIFOLTIRNGIPLFQVVDNDSRSSRSLILGATM
					KDSDSEKLDTSVQCTFYFSVNEVKKKSGSWLIHGHEKCGFVNLIGOMLGNSHS
					VDISEOKLDESVLDFEVSQVGRKEVAWIKKRPVGSITSFETSILIPGCVHS
					FPEKGAPSLISRWSGGCLDCGGVDGCKHLVLSNKTVLHKTNQSLTDFEVSQED
					SSPALAMTELKTIYRVEFGSVSPLOAFVVCVTLTCAKAKTTGKSSSPMAPPLSP
					VGRV"
					8083..8213
					/gene="F7J8_10"

JOURNAL MEDLINE REFERENCE AUTHORS TITLE JOURNAL	physically assigned pl clones DNA Res. 4 (3), 215-230 (1997) 97471969 2 (bases 1 to 83599) Direct Submission Submitted (02-JUL-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934) Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/eqd_graph.cgi?c=MM12 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlin.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MBK20.
FEATURES source	Location/Qualifiers 1. .83599 /organism="Arabidopsis thaliana" /strain="Columbia" /db_xref="taxon:3702" /chromosome="5" /clone="MM12" /clone_lib="Mitsui pl" join(2023..2363,2483..2573,2733..2804,2904..3005, 3123..3155,3277..3366,3503..3553,3630..3711,3915..4009, 4090..4161,4233..4349,4501..4569,4679..4736,4833..4954) /note="contains similarity to unknown protein gb AA45601.1 gene_id:MM12.2" /codon_start=1 /evidence=not_experimental /protein_id="BAB09942.1" /db_xref="GI:10176712" /translation="MSLVDISGAYSLVPLPPPPPLMRRAPLPPPPPLMRRAPP PPPPPLMRRAPP ECDDIEQILCAIELSVSEITFSLGAKPKPKPPEKVPPLDILRRATNTEIRMLL NIRLPMIAAAMADESLDDFDQIENLNFPTKEDMKFLTYTGDKNGCEQLFOYL QEVVYPRVESKLRFVSKFQFQGTQITKLTGLNVAVNSACEEITRSOKLDMENILC LGNLNGQTRGAVGFRDLSLLISETRADNSKMTLMHYLCKVLASKDLDLDFHD LESLEASQTLASAEITQAITKLEKLEKLELTASETDGVPVQVSKRLKLETSNNE TQVATVVALYIPARGNAEALAHFHYHYPFQVATLLSFIRLFKAHEENKVAQELE KKAKEAEKMTK" join(6389..6616,6703..7246,7345..8423) /note="gb AAP67766.1 gene_id:MM12.3 similar to unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB09943.1" /db_xref="GI:10176713" /translation="MOQNSFFLGVNRDITDARTASGWSLCPVLDDITRYGVNVHDD TFSEFLSDLYLRQVKALRHTMLVHESVFESQICELHRLYRKOKELMMETRHNA LYLNSGLPRTTHWMSNSIAYOTRNLPHEENISLVDNKKVEKEKXKVLDELPLP FEYSDLEEVHEAQNTLEQSLQMSLDQSKSKQLDNLNEPAKTEEHSDFYFNQFL SSVYISIEGNETKNEGSLVKGNSGINEAQGVKCRDIEDYDLNLSPPSSSEELTI VKFETKPOESSVSLHGHGUDQGVVQELPCLTSTLLDKRYKSLMTRSRNRK VKHCPNKNKTFKGSDDHPKSGQASSESQSNQTSMEKSSSLSEAKSKAGTSLGNKR HCKPQKSSKGQVAKRGRVAKRRSRISLVAEGNYQVQVSAEALVIDSRSGRET
CDS	ACITSLSKNLLNFPADISSVAEDNKIDTLTQLTEKKLEEHETNLRNSSDNIAAVSSI VLKQRRSRARQKRCQDDQHVDTFSENEANDDQVIGKLMKMEASDLKWNCFPKNK SKSPKPKVDFTSSGEETGVDGKALRRGSRIPAADFKHMIIINQVY complement(join(9033..9215,9319..9492,9590..9725, 10412..10494,10570..10728,10817..10887,10981..11557)) /note="gene_id:MM12.4" /codon_start=1 /evidence=not_experimental /product="dimethylaniline monooxygenase-like protein" /protein_id="BAB09944.1" /db_xref="GI:10176714" /translation="MVTFTSEASRSRKKVCVIGAGPAGLVASARELKEKHVVVLEQ NEDVGQWFTQPNVEEDPLGRSGSINGELKHVSSISLRLTSPREIMGYSDFPFL AKGRDMRPFEGHKELWLYLKFSEAGFLREIRFNVRVFEKEEEDVDKWKVRS REKFSKPMEEIFDVAVATGSHPLRPSIKGMDSMKRQIHSHVYRVPDPFRNEVY VVGNSMGODISMELVAKESHLAKTLDISGSLSKVISHKPHLLHIQIESLEDD KGVIFDGSWVADTILYCTGYKYPFLESKGRIEVDDRVGFLPFHTPPCGLSPSL SFVGIDPKLIGFPFFEAQAKWIAQVLSGKSLSPSQMLQSVDEFRSRLAGVPKHN THDIADFTCDKYADYVGFPLEDRKLLCLSLNNSQENLETYRDSWDDHQLLEAL QSSHTFNFS" complement(join(13728..13829,13913..14094,14202..14281, 14411..14544,14670..14772,14856..15043,15118..15249, 15690..15767,15849..15989,16131..16292,16451..16549, 16832..16977,16814..16924,17022..17123,17251..17450, 17559..17624,17798..17969,18370..18429,18661..18751, 18958..19031,19255..19430,19715..19847,19928..20287, 20456..20800)) /note="contains similarity to HepA-related protein Harp gene_id:MM12.5" /codon_start=1 /evidence=not_experimental /protein_id="BAB09945.1" /db_xref="GI:10176715" /translation="MEVSEBQRRRDEATLAFAEKRSQSDSENPOQODFLAKCR KLDGNDVCPQEGYRNYEVRVLTSPVEKFRVLETCSPDSFSTVPLQGRCPCE QECRLQRLKILSDAIPLHYTQNDGGKAGYKIDYKNVSGCLSKRSKSYEVGVPWKT LAVVEQLSQTISGKWPQCLPEHYTEKVEQLIETLPKLVNALLPQLQDLRFLRR GGCFIADENGLKTQAIAGIAGVCSIGSILVYCPAVLRFWAEELERLWLPSCLPSPD VHLVFGHDGMPAYLPKPKVIVSYKMLQHLRTMLERWALLIVDSHHLRCSKKRS DPEIKTIVLDVAEKVKHIIILSGTSPSVRGLGKDKYEAKTYCEGLVGRQGGKIF QDFSKTRLEILLNLTVMIRLKHLLTQLPKPRQIVTILKKSDIALAMAVIS EAKKQDGAETVEKSHEDONARSGNEAGHVNAENSDGNSDKNQLCGKLSYQOL GIAKLSAFREWLHPLLSGLDTPTEIDGRSTRKMYVFAHHKHLVDGQEFICDKG IGFVRDILPRDLQVLSQFSFSEVKIATIGVEAGVGGLDFSAQNVVFLFELPCT PSLLQAEDRAHRRGQTSVAVNYIFCAKDTMDESNNQNLKLRHSITDQDGTGKT EIBLGAASIFKPAEESEREVLQGPSESTVADKIVESCDDPDTVDYDNTDLK DMTSQLLEPCFVENGSGMRSGTISLTMLAOENHHPKNIADGLVKEVD SSSIFPLDLSLREFVQNTGRHLYSCIPCKDPRPHFONFRPEEASNPSSQGTTK EKNPESITDDPVHVALIERPKWKSILRPIEKRLKLGKLPQLPLSLSLSTSHN SEGLLRGSKRRNTPFSEISIPLENAAVWIKVLRSGHQRKEITYQAWMSNEPLCK LCQKPCAGYNAKEPEYLEDLFCELACYEDYRTSSRYIRQIEHGTCTNCKDCHLV KLRLPLLEKRTYINKAPELFARKNLTLVNDPTEGNWADHIIPIYQGGECR LENMRTLCAHADVTAAOACERKLIRSKARKOLKNTLNELRNPNKQKLSADENTKE TDSATNEEDELMEVPGSAVSDQKINHDS" complement(21177..22862) /note="gene_id:MM12.6 unknown protein" /codon_start=1 /evidence=not_experimental /protein_id="BAB09946.1" /db_xref="GI:10176716" /translation="MSEENISQVDDKCLDADNPMEISVETIPVDEEVGGDDVLHP AAETNHLAKRYSTRVLDKGTQVTRYGNQTSHTDLCKHGRREDLVKPKW VAKKNVDSGLDKGETLRKSLGNVSKPDKSQAARAAAGVAVSCDGLHVKKST KSTSPSVPVRTVKTNIIVNKAISRTSONKSPKEDLSKLNKKEGKIIVPVRCDDV LEKTDLVKVKVRSISENKSSEDTLNKEKAKIDEVPCDDVLEKSLDAQKVRISE NNKSKERLANKANENINDEPRDDDAVEKTLIVVVEVERKKKMKSTVKISSET QQSSSEKRIINSTCKSLSLPPPPSEVVTGSDPPRIQTTSKRTSLPEKKQSSGA NLVNPKPESIRKPRIGLTKTQPMNFKKGKLEKPEPSTTTSTIKFKVY VQPKPLTSDVNNKKSLDKREGVGRINGEGRKRWLRKHVEVKKKQTLFLNVI EETVNFVEKSKVAVGAFVTLSDQNNRTSOKKKIQSKTSQSSQVVEGC" join(26484..26708,26805..26841,26937..27120,27219..27367, 27442..27678,27777..27966,28070..28195,28306..28405,
CDS	physically assigned pl clones DNA Res. 4 (3), 215-230 (1997) 97471969 2 (bases 1 to 83599) Direct Submission Submitted (02-JUL-1997) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research, 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel: 81-438-52-3935, Fax: 81-438-52-3934) Address for correspondence: kaos@kazusa.or.jp For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/eqd_graph.cgi?c=MM12 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlin.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MBK20.

	/gene="F13G24.30"	
	/number=3	
Intron	11716..11813	
	/gene="F13G24.30"	
	/number=3	
exon	11814..11962	
	/gene="F13G24.30"	
	/number=4	
Intron	11963..12036	
	/gene="F13G24.30"	
	/number=4	
exon	12037..12273	
	/gene="F13G24.30"	
	/number=5	
Intron	12274..12371	
	/gene="F13G24.30"	
	/number=5	
exon	12372..12561	
	/gene="F13G24.30"	
	/number=6	
Intron	12562..12664	
	/gene="F13G24.30"	
Query Match		
Best Local Similarity 7.4%; Score 119.2; DB 13; Length 88095;		
Matches 476; Conservative 0; Mismatches 463; Indels 29; Gaps		
Qy	507 AGCTGTGCAGCTCACCRAGCTCAAGGACGGCGTCCACCATGGGATTAGCATTTAAACCATGCC	566
Db	16508 ATCTTTTCAGTAACAAGAATTTGGTAGTGGAATGTTCAATGGATTTTCATTTGAAATCAGTCCACG	16567
Qy	567 TGTGCTGGATGGTACTTTCGACGCTGGCACATTTATGACCTCGTGGTCCGAGCTTTGCTGTG-	625
Db	16568 TGTGGAGATGGAGTTCTTGTGGCATTTTTTCAACTCTTTTGTGCGAGATTTTCAATGC	16627
Qy	626 ---GGTCCACCTCAATTTCTCTCCACCATTCCTTTGAACGAACCAAGGCTCGTAACACTC	682
Db	16628 ACAAGAAACCGATAACTTGTCTCTCAAGAAATCCTCCAGTCCCTTTCCAGTTCGTTTCCCTAA	16687
Qy	683 GAGTCAAGCTCAACTCTCTCAACCATCAGATGCACCCGNACATGCTAAGTCACGAACCA	742
Db	16588 AGGTACGGT---CCTGTGTACAACCTTCCTTTCACTCCTCATGTATGAATTCATTAGCCG	16744
Qy	743 ACGGTGATGTCCC GCCCAAACGTAGACCCACACTCTTCGCGAAAAGAGTATTTCAAGTTCTCCG	802
Db	16745 ATTGAATCACCGT-----TTGAAGGAGAGAATATTCATTTCTCGT	16788
Qy	803 AGTTAGCAATTGACAAATCAAGTCAACAGTCAATGCCAATCAGGAGAGACGCCATTTCT	862
Db	16789 CTGAACAATAAACATCATTTGAATCAAAAGGCTAATGAAGAAAGCAGACAACAACAATTT	16848
Qy	863 CCACATTTCCAATCACTCTCCGACACAGTGTGGCTAGCCGTCACAGCTCGCGGCCCAACTCA	922
Db	16849 CTTCTTTTCAAGCGTTAGCCGCATTTATGTGGAGATGCAATTTACAAGACAAGAACAATTTAC	16908
Qy	923 AGCCCGAGGACTACACTGTGTACACTGTGTGTGCTGATTGTCAGGAAAAAGGTTGATCTCTC	982
Db	16909 CTTATGATCATGAATTCGTTGTAGCCTTCGCGCTAACAAANTGGAACAANAATTGGATCCAC	16968
Qy	983 CAATGCCAGAAAGTTACTTCGGCAACCTTAATTCAGGCCAATTTTTCACAGTGACCG---	1039
Db	16969 CATTTGCTTTGAGCTATCTTGTTAAATCTCTTTCTGCTGTAAAGTCCAAAGCCGTTACAT	17028
Qy	1040 CAGGTTTCTTACTAGCAAGCCCGATCGAGTTTCGCTGGTGGGATGATACAACAAGCGATCG	1099
Db	17029 CCGGTGAGCTACTAGAAAATGATTTGGGATGGGCTGCTTTGAAAATGCATGAAGCTGTGA	17088
Qy	1100 TGAAGCATGACGCTAAGGCCATTGATGAAGAAACAAGGAGTGGGAGAGCAACCCGAAGA	1159
Db	17089 TTGGAAACACTAGTGAAGTGGTCTCGGAGACAAATCAAAAACCTGGTTGAAATCATCTTATG	17148
Qy	1160 TCCTTTCAGT---ACAAMGATGCTGGAGTGAACCTGTGTTGCTGCTTTGGAAAGTTTCGCCAAGGT	1216

[illegible]

Genes were identified by a combination of three methods: Gene prediction programs including GRAIL (<ftp://arthur.epm.cornell.gov/pub/graill>), Genefinder (Phil Green, University of Washington), Genscan (Chris Burge, <http://genomic.stanford.edu/GENSCANW.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins.

Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the CSHL/WashU/ABI consortium for sequencing BAC clones F623, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khalak, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tigr.org.

```

FEATURES
    source
        1. .110531
            Location/Qualifiers
                /organism="Arabidopsis thaliana"
                /cultivar="Columbia"
                /db_xref="taxon:3702"
                /chromosome="II"
    misc_feature
        <1. .1483
            /note="Sequence from clone T517"
    repeat_region
        660. .746
            /rpt_family="POLY A"
    mRNA
        join(<770. .853,968. .1059,1190. .1278,1443. .1674,1811. .2363,
2453. .2556,2790. .3035,3146. .3272,3336. .3579,3664. .>4244)
        /gene="At2g39850"
    gene
        <770. .>4244
            /gene="At2g39850"
            /note="T517.15"
    CDS
        join(770. .853,968. .1059,1190. .1278,1443. .1674,1811. .2363,
2453. .2556,2790. .3035,3146. .3272,3336. .3579,3664. .4244)
        /gene="At2g39850"
        /codon_start=1
        /product="putative serine protease"
        /protein_id="AAB95271.2"
        /db_xref="GI:6598792"
        /translation="MATAFHTFLQLLFFFAFAEANDSRKTYLVQMKVGGCHYCSS
SQHQLLEGLVDDSYKESFTGFSASLTPRERQKLSKTTVSSRRLEVSRNKL
KLOTRSWDFMNTLKAERNPENESDLVAVIDSGIWPYSELFGSDSPPPGWENKCE
NITCNKIVGARSYPKERYKWEKSVIDVTGHTGTHVASIVAGRKVEKAGYFGLAE
GTRGVPNAKIAVYTCWRVIRKNGREDSVCREDNLLKAIDDAIADKVDIIISVSQF
QFTPLQDKVSWAFLRALKNGILTSAAAGVYANNKFPYTVANGAPWMTVAASLKDR
IFETKLEGEDKPIIVDPIINTFETQDSYPLLNKAPPSTRKRELIAERNGYSIL
SDERDKGDFVEFAQINLLDEAIKEREGKALVIGKSDIFNESIKLQFPFIASIFL
DEQKGLKLDYKKDQSKERLAKIHKTEEPREEGWPTVAHLSSRGPNCSDFLANIL
KSHMNNCFQDIAAPGLDIAGPENVKLSRDPANDYRHLRFNIMSGTSMACPHT
GLALYSKFRKWPSPSAKSLMTTSSMTDDNEFAYGSGHLNATKVRDPLGVYETHY
QVDIYLCKIAGYTEKLRSVHGSDKIDCSKTEIDHDADLANPTMTARVPLDTPPKK
VFHRTYTVNDGGETYLREINRGDKRDFDEIIVDPOLKSELGETKFTFTVTGISK
RNKNKRAFWTRNTLWLTWTEKDSQVRSPIVIYSINGPKACM"
        /note="Sequence from clone T28M21"
        complement(9924. .9995)
    tRNA
        /gene="At2g39860"
        /product="tRNA-Ala"
        anticodon=(Pos:9960. .9962,aa:Ala)
    gene
        complement(9924. .9995)
        /gene="At2g39860"
        /note="T28M21.2"
    repeat_region
        complement(11732. .11793)
        /rpt_family="{GAAA}n"
    mRNA
        join(13781. .14113,14198. .14723,14813. .14914,15003. .>15094)
        /gene="At2g39870"
    gene
        13781. .>15094

```

```

/gene="At2g39870"
/note="T28M21.3"
join(13841. .14113,14198. .14723,14813. .14914,15003. .15094)
/gene="At2g39870"
/note="unknown protein"
/codon_start=1
/protein_id="AAB95272.1"
/db_xref="GI:2088640"
/translation="MSRLLPHLVSNLPTMEKTTSSSELTPTRLGFRNFPYEFDPSPF
SPGTSPEDSSTEDSEDDFLAGLTRRLAPSTORLSPFKSEKQVAATSPQ
STLUGGSFNSGSRSDLPSPAPTSFRRDNAMDVISAAGEVAERKLKGSYEPHLL
PLQTPESLLRRQNAIIHAELQHORIEMMLCSAQSRFKLSENRIPRRVNEEGLFN
PRVRRNPTWLPQQAAPAKRPSAGTGVFLPRYPSPAAPSDSLKTPVNTPAMLOPK
VKQNLNDFEFTNIVGPRRSQFDYECMLARSTVLARQGNFRAVSGGGLPQDWMY"
complement(15255. .15311)
/rpt_family="(TA)n"
/rpt_start=15812
/rpt_family="(GAAA)n"
complement(join(<18089. .18873,19305. .>19623))
/gene="At2g39880"
complement(<18089. .>19623)
/gene="At2g39880"
/note="T28M21.4"
complement(join(18089. .18873,19305. .19623))
/gene="At2g39880"
/codon_start=1
/product="putative MYB family transcription factor"
/protein_id="AAB95273.1"
/db_xref="GI:2088641"
/translation="MNGEISRPPELISRNPKCSFENAIHKAVEAEALAKSDANG
GKSKVGPWLPEQDEALTRLVKMGCPRRNNLISRGIPRSGKSCRLRWCQLDPILR
KPSDEEHMIMSAQAVLGNKWSVIAKLLPGRDINAIKNHNLSRLRKPRAEOWKIPLL
MSNTEIVQLYPSMVRISNAKPEHLPOBEETCVLSDDKMDDEAKPEPRONSKTGV
YRVARMGATSVCKPGYMAPEGLPVQASRPSDLAGFLQSLCYDPIIPSKRGCCN
HQDTSSSVLGSSEFVDEEHSAAELDKELISINDLANTAWIRSGKEAQSLKAD
DQFRREYAHKFSGMVNNVGSQMVQRDLRLS"
complement(20519. .20562)
/rpt_family="(TA)n"
/rpt_start=22037
/rpt_family="(TAAA)n"
22753. .22816
join(<26166. .26265,26637. .26861,26940. .27034,27116. .27305,
27400. .27524,27669. .27884,27969. .28451)
/gene="At2g39890"
<26166. .28451
/note="T28M21.5"
join(26166. .26265,26637. .26861,26940. .27034,27116. .27305,
27400. .27524,27669. .27884,27969. .28346)
/gene="At2g39890"
/codon_start=1
/product="proline transporter 1"
/protein_id="AAB95274.1"
/db_xref="GI:2088642"
/translation="MTATEAKNRKINVGDDVDVIEIPDTAHOISSDSWFOAVFLT
TGNSAVLVGSGTLMVPLGNIGVYVCLLIATLSLVANTLIJAKLHFEFGRRHRYRD
LAGTYGRKAYHLTGLQYVNLFWINGCFILAGSALKAVYVLFRRDHTMKLPHFIAL
ALGICAFIAGPHLSALGVMLGVSTLSLIYIVAVLSVRDGTGKDYEIQGSLS
LSKLTITGAANLVAFNTGMLPEIQATVQRPVKNMKAIFYQFTAGVLPYAVTF
ISGYWAYSSTYLLNSVNGPLWYKALANYSAILQSVLSIHFAISFYEMDKYGIK
GNPFAIKNLLFRIMARGYIAVSTLISALLPEFLGDMSLTGAVSTFPLTFLANHYI
KAKNNKLNAMQKLWHLNVVFFSLMSVAAAIAAVRLAVDSKNFHFVADL"
join(28838. .28973,29224. .29414,29559. .29658,29750. .29787,
29872. .29961,30172. .>30411)
/gene="At2g39900"
28838. .>30411
/gene="At2g39900"
/note="T28M21.6"
join(29280. .29414,29559. .29658,29750. .29787,29872. .29961,
30172. .30411)
/gene="At2g39900"
/codon_start=1

```

/product="putative LIM-domain protein"
/protein_id="AAB95275.1"
/db_xref="GI:2088643"
/translation="MSFTGTQOKRACRKYVPVLLSADGISYHKACFKSHKSRLL
QUNSYSSMEGVVYCHPEQLFKESGFSKFNQSPAKPLTDKPTPELNRTPSLAGME
GSTQDKCATCTKTIVYPIEKVTVESQYHKSCFKSHGGCPISPSNYAALLEGILYCKHH

Query Match	6.6%;	Score	106.2;	DB	12;	Length	110531;
Best Local Similarity	49.9%;	Pred. No.	1.7e-14;				
Matches	267;	Conservative	0;	Mismatches	268;	Indels	0;
Gaps	0;						
QY	860	TTCTCCACATTC	CAATCACTCTCGGCACACG	TGTGGCTAGCCGTCACACG	TGCGCGCAAC	919	
Db	59140	TCCTCGCTTT	CAATCTCTATGCGCGT	TGCTTGGCGTGC	GATCACACGTCGCGAGGAAGC	59081	
QY	920	TCAAGCCCGAGG	ACTACACTGTGTCTGCTG	TATTCAGAGAAAGG	TTGATC	979	
Db	59080	TTCCGAGTT	TGGAAGACGCTACG	TTCGTATGGCTGTGAAT	TGTCTCACCGTTTGAGTC	59021	
QY	980	CTCCAATGCC	AGAAAGTTACTTCGGCAACT	TAATTCAGGCAATTTT	TCACAGTGACCGCGG	1039	
Db	59020	CTAAACTGAT	CCGGAGTATTTTCGGAAC	CGGATTCGAAGCGTTCT	CTACGTTTGCACACTG	58961	
QY	1040	CAGGTTTGT	TACTAGCAAGCCCGATCG	AGTTCGCTGGTGATGAT	TACAACAGCGATCG	1099	
Db	58960	CGCAGAGG	TTTTGTCCAGTGATCTT	AAATGGTGGCTGATCAG	CTTAACCCAGAGTGTG	58901	
QY	1100	TGAAGCATG	CGCTAGGCCATTTGAT	GAAAGAAACAGGAG	TGGGAGACCAACCCGAAGA	1159	
Db	58900	CGGCTCAT	CAAGACGGGAGATTCG	TAGCGTTGTGGCGGAT	TGGGAAGCGAATCCCAGGT	58841	
QY	1160	TCCTTCAGT	CAAAAGATGTGGAGT	GAACCTGTCTGTGGTGA	AGTTCGCCAAGGTTC	1219	
Db	58840	GTTTTCTCT	TGGAATGCGCAGGAG	CTTCGGTTACTATGGA	AGCTCTCCGAGGTTTC	58781	
QY	1220	AGTTTACG	ACGTGGATTTGGATGGG	AAAGCCAGAGAGTGT	GAGAGTGGTTCGAACA	1279	
Db	58780	CGATGATG	ATAATGATTTTCGGGT	GGGAAGCCGCTGGCGG	TTAGAAAGTGGACGTCGA	58721	
QY	1280	ATAGGTTT	GTGATGAATGGTGTAT	TTTGTACCAAGGCAAA	ATGAGGAGNAGCATTTGATG	1339	
Db	58720	ATAAGTTT	GTGATGAATGATTTTC	GGCGGTTTCCGGGGAG	GGAAGAAACGGGACTGTTGAT	58661	
QY	1340	TGGAGATT	TAGTTTGGAAAGCAAT	TGCTATGGAGAGGT	TGGAGAAGATAAAGAGTT	1394	
Db	58660	TGGAGTGG	TTTTATCCCGGACACT	ATGCTCGGATGGAAT	CTGACGCTGAGTT	58606	

RESULT	9
LOCUS	AP002842/c
DEFINITION	Oryza sativa genomic DNA, chromosome 6, BAC clone:OSJNBA0075G19, complete sequence.
ACCESSION	AP002842
VERSION	AP002842.1
KEYWORDS	HTG.
SOURCE	Oryza sativa (cultivar:Nipponbare) DNA, clone:OSJNBA0075G19.
ORGANISM	Oryza sativa
REFERENCE	1 (sites)
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Oryza sativa nipponbare(GA3) genomic DNA, chromosome 6, BAC clone:OSJNBA0075G19
JOURNAL	Published Only in DataBase (2000) In press
REFERENCE	2 (bases 1 to 88318)
AUTHORS	Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE	Direct Submission
JOURNAL	Submitted (09-AUG-2000) Takuji Sasaki, National Institute of Agrobiological Resources, Rice Genome Research Program; Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
	(E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/.

COMMENT		Tel:81-298-38-7441, Fax:81-298-38-7468) The orientation of the sequence is from -21M13 to M13rev of the BAC clone.									
FEATURES		Location/Qualifiers									
source		1..88318 /organism="Oryza sativa" /cultivar="Nipponbare" /db_xref="taxon:4530" /chromosome="6" /clone="OSJNBA0075G19"									
BASE COUNT		24979 a 19394 c 19436 g 24509 t									
ORIGIN											
Query Match		6.4%; Score 102.8; DB 13; Length 88318;									
Best Local Similarity		46.1%; Pred. No. 1e-13;									
Matches		455; Conservative 0; Mismatches 522; Indels 9; Gaps 3;									
QY	418	CTTACTGATGAAGAAGCACCACCAAAATTTGCGAGGACTTGATTCTTGTATAAATAATCTTG	477								
Db	86254	CCTCCGACGCCACGTCGCCGCCCTCTCCATCTCTTCCCGCTCGGGAGCGCGTC	86195								
QY	478	AATTTGGAAGGCTTCATCGCCCTCTTCTAGCTGTGCAGCTCACCAAGCTCAAGGACGCGG	537								
Db	86194	AACACGAGGCCACCATCTCCCTCTTCTGTCGTCAGGTCACCACTCTGTTGACGGC	86135								
QY	538	CTCACCATGGGATTAGCATTTAACCATGCTGTGCTGGATGGTACTTTGACGCTGCACCTTT	597								
Db	86134	GTCTTCTCTCTTCTGCTCTACAAACCACGGCTCTCCGATGGAAACCGCATTCGGAGCTTC	86075								
QY	598	ATGACCTCGTGGTCGAGCTTGTGTGGTTCACCTCAATTTCTG-----TCCCAACCAT	652								
Db	86074	CTCAATGCGTGGTTCGAGATTGCGCGCGCCAGCTGCCTCGCCGCGCCACCTCGCCCCCA	86015								
QY	653	TCCTTTGAAGAACCAAGGCTGCTAACTCTCGAGTCAAGCTCAACCTCTCTCAACCATCAG	712								
Db	86014	CCCTTGTTCACCGTTTGGTCTCCCTCTCCGGAGACGAGCCCGGTGCTGCTGCCGTAC	85955								
QY	713	ATGCACCCGAACATGCTAAGTCAGAACCAACGCTGATGTCCCGGCCAACGTAGACCCAC	772								
Db	85954	GCCGACTGTCGGAGCTGATCAGAGGCTGCAGCGCGCTGTGTGCGAGCGGATGCTA	85895								
QY	773	CTCTTCGCGAAAGAGTATTTCAAGTTCTCCGAGTTAGC-AATTGACAAAAATCAAGTCAACA	831								
Db	85894	CACCTTCTCATCGGAGTCGCTGTGGCGCTCAAGGAGCGGGCTCGCCAAGAGCTCCTGGCG	85835								
QY	832	GTCATGCCAACCTCAGGAGACACGCCATTTCTCCACATTTCCAATCTACCTCTCCGACACAGT	891								
Db	85834	GCCGGGACACGGGGCGCGCGCTGACGAGGTTCCAGGGCTGAGCTCGTCTGCTG	85775								
QY	892	TGGCTAGCCGTACACGCTGCGCGCCAACTCAAGCCCGAGGACTACACTGTGTACACTGTG	951								
Db	85774	TGAGGTGCATACCCCGCGCAGCGGCTGCGCGCGGAGCAACAGCATGTGTCGCGCGG	85715								
QY	952	TTTGTGATTTCGAGGAAGAAGGTTGATCTTCCAAATGCCAAGAAAGTTACTTTGCGCAACCTA	1011								
Db	85714	GCCATCAACAACCGCGCGCTCTCCAGCGCGCTGCGCCAGGAGTACTTTGCGCAACAGC	85655								
QY	1012	ATTTCAGGCAATTTTTCACAGTGCACG---CGGACGGTTTGTACTAGCAAGCCGATCGAG	1068								
Db	85654	ATCTACGCCATCAGCACGGACAAGGTCAGCGCTGCGAGCTGTTGGAGCGCGGACGCGG	85595								
QY	1069	TTCGCTGGTGGGATGATACAACAAGCGATCGTGAAGCATGACGCTAAGGCCATTTGATGA	1128								
Db	85594	TGGCGCGCGCGCGGTGGGCTGGGCTGGCGCGCACACGCGACGCGGACATCAGGGCG	85535								
QY	1129	AGAAACAAGGAGTGGGAGAGCAACCCGAAGATCTTTCAGTACAAAAGATGCTGGAGTGAAC	1188								
Db	85534	CGGTTGCGCGGTGGGAGGTTCAAGCCGATATATACACGGCGAGGTACTTTCGACCCGCTG	85475								
QY	1189	TGTGTTGCTGTTGGAAGTTCCCAAGGTTCAAGGTTTACGACGTTGGATTTTGGATGGGA	1248								
Db	85474	GGGTTGATGATGGGAGGCTCCCGGGTTTCACATGTACAGGCTCGCACTTTGGGTGGGG	85415								

exon	AYFSAVAIVILLTLQSFYAVYAYKPN SKL" complement(2238. .2668) /number=1	gene	/number=4 12843. .13704 /gene="F11C1_40"
Intron	complement(2669. .2806) /number=1	CDS	complement(join(12843. .12868,13035. .13704)) /note="similarity to predicted protein, Arabidopsis thaliana"
exon	complement(2807. .3330) /number=2		/codon_start=1 /product="putative protein" /protein_id="CAB62299.1"
Intron	complement(3331. .3448) /number=2		/db_xref="GI:6523031"
exon	complement(3449. .4119) /number=3	exon	complement(12843. .12868) /number=1
gene	/gene="F11C1_20"	intron	complement(12869. .13034) /number=1
CDS	complement(join(5324. .5760,6099. .6562,6678. .7543)) /note="similarity to various predicted genes, Arabidopsis thaliana and Oryza sativa"	exon	complement(13035. .13704) /number=2
		exon	complement(14718. .14858) /number=1
		gene	/gene="F11C1_50"
		CDS	complement(join(14718. .14858,14985. .15023,15097. .15177,15266. .15312,15360. .15461,15544. .15660,15743. .15804,15889. .15935,16035. .16098,16167. .16285,16375. .16490,16576. .16609)) /note="similarity to SRG1 protein, Arabidopsis thaliana, PIR:S44261"
exon	complement(5324. .5760) /number=1		/codon_start=1 /product="flavonol synthase-like protein"
Intron	complement(5761. .6098) /number=1	exon	/protein_id="CAB62300.1"
exon	complement(6099. .6562) /number=2	intron	/db_xref="GI:6523032"
Intron	complement(6563. .6677) /number=2	exon	/translation="MATDFKSLPVIDISRLLLKCDPDMADVGVAEVVQQLDKACRDAGFFYVIGHGISIEDVINKREITREFFKLPYEELKIKMTPAAGYRGVORIGENVTKGIPDIHEAIDGKYGDIGKVMGPNQWPNPQEFKELMEEYIKLCTGDLRKLRLGSLA
exon	complement(6678. .7543) /number=3	intron	LAGSPYEFEGKMGADPPFWVRLIGYPAQVTNLFRSSAYNDMFQVLTUTMTVTKFNPVEVKSSLLTLNQDDDKTALQVRNLGEMISAIPIPGSFVCNIGDKMLKILSNGVYESTLHR
gene	/gene="F11C1_30"		TNFDVAVPELDICKOKYPGGRGQCQVFRAYVEGHLVSKVQTNFAM"
exon	complement(9545. .9615) /number=1		/number=1
CDS	complement(join(9545. .9615,9706. .9978,10386. .10813,10931. .11350)) /note="similarity to various predicted proteins, Arabidopsis thaliana and Oryza sativa"	exon	complement(14985. .15023) /number=2
		intron	complement(15024. .15096) /number=2
		exon	complement(15097. .15177) /number=3
		intron	complement(15178. .15265) /number=3
		exon	complement(15266. .15312) /number=4
		intron	complement(15313. .15359) /number=4
		exon	complement(15360. .15461) /number=5
		intron	complement(15462. .15543) /number=5
		exon	complement(15544. .15660) /number=6
		intron	complement(15661. .15742) /number=6
		exon	complement(15743. .15804) /number=7
		intron	complement(15805. .15888) /number=7
		exon	complement(15889. .15935) /number=8

http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=k21h1
Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, <http://combio.ornl.gov/Grail-1.3/>), GENSCAN (Chris Burge, MIT, <http://CCR-081.mit.edu/GENSCAN.html>), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://grain1.zoool.iastate.edu/cgi-bin/sp.cgi>).
Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>).
This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K8A10 and the 3' clone is K3G17.

FEATURES

source

1. .74342
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone="K21h1"
/clone_lib="Mitsui TAG"
complement(join(683. .778,1044. .1247,1325. .1959,2053. .2283, 2364. .2661))
/note="gene_id:K21H1.1
unknown protein"
/codon_start=1
/evidence="not_experimental"
/protein_id="BAB10939.1"
/db_xref="GI:10177592"
/translation="MNMSDDDEPRGYILRPEELRPWELVRLFFSGDIEKPRSVSDS ETEHRSFRRWIFVSLVLLKLLRFFSKLLALVGSALFSLNFSNNSFGFLRGEV VQPTSENYSQFICHLDTRVSLDMLNREDEKYYAALSIMASKIAYENAAIRKHVV ENHNMKVLGVLYNEXOEKETTOAFIMSTDETTTRNGOETTUVVAFRCGTELFNSE DMCSDPDTWELPNIHNGHGFMAKGLGNCCNWPKEPLSPDRKSPPLAYISRDSL KTLIAQNKTFLVLTGSLGGALAILFTAVLVIHHTELLERIQGVYITGQPRVGDGSG FGEFMEKLEKRIKYYRFVYNNDIVPRPYDDKDLKFKHGTCTIYDQNYOAKYMR ESDNEFFLLRGIMKWSAILEFIRSFITVAEKSEYSEGWLLKGALGIGTIVGVSN HTPOYKNSRVVSYNGYEDVPRKDETALKRVIFPTG"
7772. .8497
/note="emb|CAB62312.1
gene_id:K21H1.2
similar to unknown protein"
/codon_start=1
/evidence="not_experimental"
/protein_id="BAB10940.1"
/db_xref="GI:10177593"
/translation="MDSDIMMMHMQEKLPEFCNPNSFFSPDINNTYPPFLFNSTHY QSDHSMTEPGFRYSGGLLTNPSSIPNTAVSYVFLDKRNNNNNNNTMAAMREMI FRIAYQTHIDPEAVKPKRRNRVSKDPQSQVAARHRRERISIRILQRLVPGCTK MDTASCKATHYVFLKQVOSLEQAVVTGGGGGGGGRVLIGGGGTAASGGGGGG GVMVGCGTCTGTHQVGNQILR"
9616. .10005
/note="gb|AA64534.1
gene_id:K21H1.3
similar to unknown protein"
/codon_start=1
/evidence="not_experimental"
/protein_id="BAB10941.1"
/db_xref="GI:10177594"
/translation="MAASNLNLLLSLTLFTLSLQSRSESLSDNPSTLILPDGFWPIS HSDEFDIDGESFVETEDDGVTDRRSLWRRTYYISYGALSANRPVPCPSCRSY YTHNCFRAGPVHPYSRGCSSTTRCR"
14222. .15256
/note="gene_id:K21H1.4"
/codon_start=1
/evidence="not_experimental"
/product="protein kinase-like protein"
/protein_id="BAB10942.1"

CDS

/db_xref="GI:10177595"
/translation="MEWIRGETIGYGTFTSVLSTRNNDGCEPPLMAVKSADSYGA ASLAKSVLNDLDDCNEIYRCFGEEDRTVENGEEMHNLFLFYASRGSLESYLKLAG ASPESTVRRHTGSVRLGRHLIHANGFAHCDLKIGNILFUDGAVKRIADFLGAKRIG LTLALNYGOIRGTPLYMAPESYNDNEYSGEDVWALGVVFMFGSKTAWSLKSNF MSLLRIGVDEVPIMPELSSEQDFLSKCFVKDKRRTAEMLLNHPFVTVDDHDL VLVKEEDFVWNNKEDVSTSPRCPEFPDMVSVSSGQSTIDSPDERVASLVTMDIPDW SVTNSWTVNR"
complement(15607. .17817)
/note="gene_id:K21H1.5"
/codon_start=1
/evidence="not_experimental"
/product="subtilisin-type protease-like"
/protein_id="BAB10943.1"
/db_xref="GI:10177596"
/translation="MGMTVILVFPFVAIVTAETSPYIIHMDLSAKPLPSPDRHSW FSTPTFSLTNKRKPIIYAYDSVHGFSAVLTNSLQRLKHKPGVSVFTKDLVPLKLT FTSPKFIENLSTGTPWVSYNGAGIVIGIDTIPWDSFHDGTVGVSVKWKACE FNSSSLCNKLLIGAKVFNKGLFANNPDRLRETRIGQYSSPYDTIGHGVHAAIAAGNHV KNASYFSAQGTASGIAPHALAIYKAAWEEGIIYSSDIAAIDQAIRGVHVHISLSLG LSFEDDDNUGFLENDPIAVASFAAIQGVFVVTSGGNDGVPVWSLNGAPNIWVG AGTIGRQFGTLTFGNRVSFSPSLFPGFEFVQFPVYIESGSVENETLANIYVCN ENINIGSKLHQIRFGAAAVLITDKLLEEQDTIKFPVAFIGSKHRETIIESYASN KNNATAKLETIKTVICTPAPEVGTSSRGPTSPQLKPDILAKPTLILSAMPSE QITGTRALPLFSGFNLLTCTMAAPHVAGVAALIKQVHPNWSPSAIKSAIMTALTLD NPLVAGCHVSTNKVNLNPLGIYDTPQDFINFLCHEAKQSRKLIITRSNSTDACKK PSLYLNYSIITAYFTSDQSSPKIFKRTLTNVGEAKRSYIVRVGLGNVVPVKLM FSEKNEKUSIYVRLESRGLQNVVIGLVSWDEDEAEFEVSCSVAVSLVQES"
join(18490. .19596,19689. .19859,19945. .20028,20117. .20191, 20278. .20342,20515. .20605,20698. .20883,21116. .21279, 21388. .21502,21780. .21956,22033. .22146,22391. .22510, 22590. .22904,22985. .23074,23461. .23579,23668. .23740, 23831. .23887,23939. .24091,24221. .24343,24422. .24519, 24598. .24643,24730. .24843,24949. .25032,25116. .25322, 25414. .25490,25698. .25821,25946. .25999,26108. .26227, 26298. .26372,26460. .26600)
/note="gene_id:K21H1.6"
/codon_start=1
/evidence="not_experimental"
/product="DNA polymerase alpha 1"
/protein_id="BAB10944.1"
/db_xref="GI:10177597"
/translation="MSGNSTETGRRSRGAEASRKDTLERLKAIRGOGIRASGGG YDRIQKPIEDTVDDDEYDALVSRREERARGPVVEDGGDGLYIDEGEEDWSPSG PESTDESDDGGFRSGRLAKKKKKEQTOQPOQVKKYNPALKAAATITGBRLSSMTSS SFKVKETDFIAQYEGILDEITAOVTPDESRRKRRRLPGTVPTIFANKKLFVSAS SMGKSESEPTSTYEGDSVMDNLMKEEDMKESVIPTSETMELIGSDIVKEDGSKNI RTTEYKSELGVKEVFTLNATIDMKEKSALSATAGKEMKGVGTENGALLGSSEGG TEFDLADGSRFFILDAYEEAFGASMTIYLFKGVKMGDTYKSCVVVKNQRCVYA IPDSIFPSSHLLMLEQEVKDSRLSPESFRGLHEMASKLNEIAQELLQLVNSFSM APVKRYAFERDPVPAGEQYVLKINYSKDRPLPEDLKGESFALLGSHTSALEHFIL KRKMGPCWLKISSFSTCSPSGSWCKFEVTVQSPKDTILVSEKVVHPVAVTAL NUKTIVNEKQNISEIVSALVCFHNKADVPMPAPERKRKSDLSLTHVYVNPPEGTY PIGWKEVDFNSKNCNVLSEIENRALLNRLLELNKSDILDLVHNSIFGDLVLL ORAQACKVYSKSMWSKIGRLKRSFMPKLGNSNYGSGATPGLMAGTLLCOTDLC SR DLLKEVSYSLDTLSKTLNDRPKETAPNDIPKMFOSKTLVELIECGETDAMLNLM PHLSVLPTLQTLNLSNGLWKTLOGARARQRIEYLLHTFHSHKFTLPDKISQRKEI KSSKRRMDYAPEDRVNVDLADLTLENDPSGSKTKKGPAYAGGLVLEPKRGLYKYV LLLDFNSLYAFSIQYENICFTIPRSEDGVPRLSPSSQTPGILPKMLHLVSRKSVKL RKMKETGLKYWELDIRQOALKLTANSGLCFNSRFYAKPLAELITLQGRDILQRT VDLVONHLNLEVIYGDTSIMHSGLDDEEVKAISKVIOEVNKKYRCLKTDGCIY KRMLLRKKKYAAVKLQPKCKPCEDIERKGVDMVRDWSLLSKEIGDCLSKILYGG SCEDVEEAHNEMLKIKEMRNGQVALEKIYITKTLTPPAAYIPDSKSVOPHVVQVALRM RORGYEFNAKDTVPYICYBOGNASSASSAGIAERARHPDEKSESRMLVDIYY LAQOILHPVYSLRCAIQGTSPERLAECGLDPSKYSKSNDSATSDPSTSLTFATSD ESKQAPTEDESSTFWLKLHCPCQOEDSTGIIISPAIANOVKROIDGVSMTYKG INVCDESKCHTTRSPNELLGERERTVCPNYPNCNGTLLRKYTEADLYKOLSVFCH ILDTQCSLEKMDVGVRIQOVERAMTKIRPAVKSAAAITRSSRDRCAVGMQLTDIYI"

CDS

complement(join(26828. .27019,27114. .27179,27258. .27323, 27414. .27485,27598. .27834))
/note="emb|CAB16798.1
gene_id:K21H1.7
similar to unknown protein"


```
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10945.1"
/db_xref="GI:10177598"
/translation="MGSDVGDRLPPSPSSDLSFLRQILSRTPTAQPSPPKSTNV
SSAETFFSVSGVSGVSGTGQDKYAFERKRGAKORNSLRKRDIAQPHNUSE
KKRSKINEMKMLQKLPINSKTDKASMLDEAIEYIKQLQLOVQTAVWNGLGNPM
RLQVPPTTRINIELEQDLNLETLLAAAPHSLEPARTSQCMCFSTATLL"
complement(join(29326..29586,29960..30082,30390..30512,
30845..30958,31036..31233))
/note="gene_id:K21H1.8
pir||T05432
similar to unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB10946.1"
/db_xref="GI:10177599"
/translation="MDPNEONSEIPDSFPMWLLFGDDMDLWDTYSPVADDDISPPD
TLVNGDGPAPLPGSLRTISWETTFQDSLQSWNDQSETTSVVEYTDYSSHGNTFT
NEETLERYWRNWLQSTNEQSETGSOEYTNASSHGFTIYEETLEQYWRNWLQSS
TNEQFETSLSEYTPNSHGDIFTYEBLLSTIDTGTGERTGLSEVIDENLRKYEK
RSDDETARCVTCQOKLDNEEVSKLGGCHDFHFGCIKNWLMVTKPCLCNREVV"
join(33235..33402,33654..33791,33873..33971,34064..34367,
34837..35225)
/note="contains similarity to MAP3K-like protein kinase
gene_id:K21H1.9"

Query Match 4.2%; Score 67; DB 12; Length 74342;
Best Local Similarity 47.0%; Pred. No. 3.2e-05;
Matches 276; Conservative 0; Mismatches 305; Indels 6; Gaps 2;

QY 817 AAATCAAGTCAACAGTCAATGCCAAGTCAAGGAGAGCGCCATTCACACATTCACATCA 876
Db 37972 AAAGCGAAAGCTAACCGCGAGGTTGACTCCGATGATCGGAAATATCGTCTCAACGG 37913

QY 877 CTCCTCCGACACGTTGGCTAGCCGTCACAGTGGCGGCCCAACTCAAGCCGAGGACTAC 936
Db 37912 GTTTCAGCGTATATGTCGCGATCGATATCAAGAACAGTGTGTGAATCCGGAAGAGT 37853

QY 937 ACTGTGTACACTGTGTTGCTGATGACAGAAAGGTTGATCTCCCAATGCCAGAAAGT 996
Db 37852 ATTCATTGCAAGTTACTGCTGGATATGAGAGAGACTAAACCCCTCCCTTGAGAAAGAG 37793

QY 997 TACTTCGGCAACCTAATTCAGGCAATTTTCACAGTGCACCGCGGAGGTTGTACTAGCA 1056
Db 37792 TGTTCGGTAACTGTCGGATTCGCGAGTGCAGGACTACTGTGGCGGAGATGCTTAC 37733

QY 1057 AGCCCGATCGAGTTCGCTGGGATGATACAAACAGCATCGTGAACCATGACCTAAG 1116
Db 37732 AATGGGCTGGTTGGGCTGCAATTCGAAATAACAAACCTGTGGGTCAACAACCAATGAA 37673

QY 1117 GCATTGTATGAAGAAACAAGAGTGGGAGAGCAACCCGAAGATCTTTCAGTACAAAGAT 1176
Db 37672 GAGTTTAGAGAGTTGCTGAGAAATTTGGGTGAAGAACCAGATATATTGAATGCTTAGGCT 37613

QY 1177 GCTGAGTGAAGTGTGTTGCTGTTGGAAGTTTCGCAAGGTTCAAGGTTTACGACGTGAT 1236
Db 37612 TTTAG---TAATTCATAACAATTCCTAGCTCTCCTCGTTCAATGTCTATGGAACCAT 37556

QY 1237 TTTGATGGGAAACCCAGAGAGTGTGAGGATGTTTTCGACACATAGTTTGTATGGAATG 1296
Db 37555 TTCGGTTGGGTTAAACCAATAGCGGTTTCGAGTGGACCGGAAATACCCCGATGGTAAA 37496

QY 1297 GTGATTTGTACAAAGGCAAAATGAGGAAGCAAGCATTTGATGTGGAGATTAGTTTGAA 1356
Db 37495 CTCATTGCTTATCCCGG---GATCGAAGAAGAAACATCGAGTTTCAACACGTGTTTATCG 37439

QY 1357 GCAATGCTATGGAGAGTTGGAGAAAGATAAAGAGTTTCCTCATGGA 1403
Db 37438 TCTAGTGTGTTGGAAAAATTTATCGACTCATGAAGAGTTTGTGAAGCA 37392
```

RESULT 14

```
AP003279/c
LOCUS
DEFINITION
Oryza sativa chromosome 1 clone P0529E05, HTG 21-FEB-2001
PROGRESS ***, in ordered pieces.
ACCESSION
AP003279.1 GI:13027309
VERSION
HTG; HTGS_PHASE2.
KEYWORDS
Oryza sativa (cultivar:Nipponbare) DNA, clone:P0529E05.
SOURCE
Oryza sativa
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae;
Oryza.
REFERENCE
1 (sites)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0529E05
Published Only in DataBase (2001) In press
2 (bases 1 to 137205)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Direct Submission
Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abrr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
NOTE: It currently consists of 1 contigs. Gaps between the contigs
are represented as runs of N. The order of the pieces is believed
to be correct as given, however the sizes of the gaps between them
are based on estimates that have provided by the submitter. This
sequence will be replaced by the finished sequence as soon as it is
available and the accession number will be preserved.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
1..137205
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="P0529E05"
BASE COUNT 38429 a 29910 c 31211 g 37505 t 150 others
ORIGIN

Query Match 3.8%; Score 60.8; DB 83; Length 137205;
Best Local Similarity 47.1%; Pred. No. 0.001;
Matches 256; Conservative 0; Mismatches 282; Indels 6; Gaps 2;

QY 860 TCTCCACATTCCAATCACTCTCCGCACACGTTGCTAGCGTACACGTCGCGGCCA-- 917
Db 19493 TCTCGTCTGTTCCAGTCACTGTGCGGCGAGATATGGCTCTCGGTGACGCGCGCGGACG 19434

QY 918 -ACTCAAGCCCGAGGACTACACTGTGTACACTGTGTGTTGCTGATTTCAGGAAAAGTTG 976
Db 19433 GTCTGGCCCGCGAGCGACGACGACGTTCCGATGGCGGTGAACCTGCCGCGCACAGGTCG 19374

QY 977 ATCTCCCAATCCAGAGAAGTTACTTTCGCAACCTTAATTCAGCAATTTTCACAGTGACCG 1036
Db 19373 GCCCGGCGACTCCCTCTGCTGTACTTCGGCAACGCCATCCAGCGCGACACACCGGCCA 19314

QY 1037 CGGCAGGTTTGTACTAGCAAGCCCGATCGAGTTTCGCTGGTGGGATGATACAACAAGCA 1096
Db 19313 CGGTGGCCGAGCTGGCGTCCCAACGACCTGCGGTGGCGCGCGCCCAAGCTCAACGCGAC 19254

QY 1097 TCGTGAAGCATGACGCTAAGGCCATTGATGAAGAAAACAAGAGTGGGAGAGCAACCGCA 1156
Db 19253 TGGCGGCGTACGACGACGCGCGGATACGCGCGCGCGCGCGTGGCAGGCGCTCCCC 19194

QY 1157 AGATCTTTCAGTACAAAGATGCTGGAGTGAACCTGTGTTGCTGTTGGAAGTTTCGCAAG 1216
Db 19193 GGTGCTTCCCGCTGGGCAACCCGCGCGGCTGATCAGTGGGAGCTCGAACCGGT 19134
```


[illegible][illegible]

Search completed: November 5, 2001, 16:46:27
Job time: 10557 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:39 ; Search time 5816.79 Seconds
(without alignments)
2635.909 Million cell updates/sec

Title: US-08-894-356C-2
Perfect score: 1622
Sequence: 1 GAACCATTCGAATCCCAATTAA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_estl1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est25:*
- 26: gb_est26:*
- 27: gb_est27:*
- 28: gb_est28:*
- 29: gb_est29:*
- 30: gb_est30:*
- 31: gb_est31:*
- 32: gb_est32:*
- 33: gb_est33:*
- 34: gb_est34:*
- 35: gb_est35:*
- 36: gb_est36:*
- 37: gb_est37:*
- 38: gb_est38:*
- 39: gb_est39:*
- 40: gb_est40:*
- 41: gb_est41:*
- 42: gb_est42:*
- 43: gb_est43:*
- 44: gb_est44:*
- 45: gb_est45:*
- 46: gb_est46:*
- 47: gb_est47:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est33:*
- 111: gb_est34:*
- 112: gb_est35:*
- 113: gb_est36:*
- 114: gb_est37:*
- 115: gb_est38:*
- 116: gb_est39:*

117: gb_est48: *
118: gb_est49: *
119: gb_est50: *
120: gb_est51: *
121: gb_est52: *
122: gb_est53: *
123: gb_est54: *
124: gb_est55: *
125: gb_est56: *
126: gb_est57: *
127: gb_est58: *
128: gb_est59: *
129: gb_est60: *
130: gb_est61: *
131: gb_est62: *
132: gb_est63: *
133: gb_est64: *
134: gb_est65: *
135: gb_est66: *
136: gb_est67: *
137: gb_est68: *
138: gb_est69: *
139: gb_est70: *
140: gb_est71: *
141: gb_est72: *
142: gb_est73: *
143: gb_est74: *
144: gb_est75: *
145: gb_est76: *
146: gb_est77: *
147: gb_est78: *
148: gb_est79: *
149: gb_est80: *
150: gb_est81: *
151: gb_est82: *
152: gb_est83: *
153: gb_est84: *
154: gb_est85: *
155: gb_est86: *
156: gb_est87: *
157: gb_est88: *
158: gb_est89: *
159: gb_est90: *
160: gb_est91: *
161: gb_est92: *
162: gb_est93: *
163: gb_est94: *
164: gb_est95: *
165: gb_est96: *
166: gb_est97: *
167: gb_est98: *
168: gb_est99: *
169: gb_est100: *
170: gb_est101: *
171: gb_est102: *
172: gb_est103: *
173: gb_est104: *
174: gb_est105: *
175: gb_est106: *
176: gb_est107: *
177: gb_est108: *
178: gb_est109: *
179: gb_est110: *
180: gb_est111: *
181: gb_est112: *
182: gb_est113: *
183: gb_est114: *
184: gb_est115: *
185: gb_est116: *
186: gb_est117: *
187: gb_est118: *
188: gb_est119: *
189: gb_est120: *

190: gb_est121: *
191: gb_est122: *
192: gb_est123: *
193: gb_est124: *
194: gb_est125: *
195: gb_est126: *
196: gb_est127: *
197: gb_est128: *
198: gb_est129: *
199: gb_est130: *
200: gb_est131: *
201: gb_est132: *
202: gb_est133: *
203: gb_est134: *
204: gb_est135: *
205: gb_est136: *
206: gb_est137: *
207: gb_est138: *
208: gb_est139: *
209: gb_est140: *
210: gb_est141: *
211: gb_est142: *
212: gb_est143: *
213: gb_est144: *
214: gb_est145: *
215: gb_est146: *
216: gb_est147: *
217: gb_est148: *
218: gb_est149: *
219: gb_est150: *
220: gb_est151: *
221: gb_est152: *
222: gb_est153: *
223: gb_est154: *
224: gb_est155: *
225: gb_est156: *
226: gb_est157: *
227: gb_est158: *
228: gb_est159: *
229: gb_est160: *
230: gb_est161: *
231: gb_est162: *
232: gb_est163: *
233: gb_est164: *
234: gb_est165: *
235: gb_est166: *
236: gb_est167: *
237: gb_est168: *
238: gb_est169: *
239: gb_est170: *
240: gb_est171: *
241: gb_est172: *
242: gb_est173: *
243: gb_est174: *
244: gb_est175: *
245: gb_est176: *
246: gb_est177: *
247: gb_est178: *
248: gb_est179: *
249: gb_est180: *
250: gb_est181: *
251: gb_est182: *
252: gb_est183: *
253: gb_est184: *
254: gb_est185: *
255: gb_est186: *
256: gb_est187: *
257: gb_est188: *
258: gb_est189: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match %	Score	Length	DB	ID	Description
1	152	9.4	583	113	AW221050	AW221050 EST297519
2	151.4	9.3	690	155	BG599447	BG599447 EST504342
3	150.6	9.3	606	113	AW221049	AW221049 EST297518
4	144.6	8.9	570	142	BE922784	BE922784 EST426553
5	140.6	8.7	763	155	BG591758	BG591758 EST499600
6	133.4	8.2	591	167	BE434257	BE434257 EST405335
7	128.6	7.9	487	118	AW616206	AW616206 EST307245
8	125.2	7.7	497	122	AW930934	AW930934 EST356777
9	124	7.6	482	144	BF096979	BF096979 EST400516
10	119.2	7.3	518	119	AW650280	AW650280 EST328734
11	114.6	7.1	564	142	BE921494	BE921494 EST425179
12	114.6	7.1	659	155	BG598154	BG598154 EST496832
13	112.4	6.9	508	173	BG097054	BG097054 EST461573
14	108.8	6.7	574	142	BE922746	BE922746 EST426515
15	103.6	6.4	403	143	BF053202	BF053202 EST438432
16	98.8	6.1	788	164	BE187619	BE187619 EST336180
17	89.4	5.5	687	111	AW102336	AW102336 sdb606.y
18	85.8	5.3	575	142	BE920468	BE920468 EST424237
19	79.8	4.9	549	21	A1490445	A1490445 EST248771
20	77.8	4.8	783	155	BG584708	BG584708 EST486469
21	75.2	4.6	451	167	BE433920	BE433920 EST404998
22	75.2	4.6	578	143	BF051101	BF051101 EST436276
23	75	4.6	542	167	BE436185	BE436185 EST407263
24	73	4.5	411	142	BE922378	BE922378 EST426147
25	72.8	4.5	456	142	BE923572	BE923572 EST426147
26	72.6	4.5	354	142	BE922950	BE922950 EST426719
27	71.6	4.4	414	142	BE921952	BE921952 EST425721
28	68.8	4.2	368	142	BE921953	BE921953 EST425722
29	67.2	4.1	603	118	AW616119	AW616119 EST296884
30	65.2	4.0	386	173	BG097680	BG097680 EST462199
31	64.6	4.0	686	120	AW776512	AW776512 EST335577
32	62.2	3.8	642	153	BG448305	BG448305 NF070A05E
33	61.8	3.8	676	153	BG455847	BG455847 NF071A12P
34	61.8	3.8	681	151	BF639071	BF639071 NF079A08P
35	59.2	3.6	694	117	AW560686	AW560686 EST315734
36	59.2	3.6	829	155	BG584678	BG584678 EST486439
37	58.8	3.6	411	142	BE921804	BE921804 EST425573
38	58.6	3.6	658	146	BF278256	BF278256 GA_EB003
39	57.8	3.6	526	105	AL386584	AL386584 MtBC35E10
40	57.8	3.6	649	151	BF643475	BF643475 NF004H05E
41	56.6	3.5	708	138	BE660849	BE660849 7-C8 Gmax
42	56.4	3.5	311	166	BE353747	BE353747 EST355090
43	55.8	3.4	307	142	BE920010	BE920010 EST423779
44	55.8	3.4	489	138	BE660848	BE660848 330 GmaxS
45	54.2	3.3	648	154	BG457806	BG457806 NF034C06P

ALIGNMENTS

RESULT 1
AW221050
LOCUS AW221050 583 bp mRNA EST 07-DEC-1999
DEFINITION EST297519 tomato fruit mature green, TAMU Lycopersicon esculentum
cdna clone cLEF3F3, mRNA sequence.
ACCESSION AW221050
VERSION AW221050.1 GI:6532734
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 583)
AUTHORS Alcala,J., Vrabalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.

Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers
1. 583
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F3"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF3: Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 142 a 140 c 117 g 184 t
ORIGIN
Query Match 9.4%; Score 152; DB 113; Length 583;
Best Local Similarity 56.4%; Pred. No. 1.3e-27;
Matches 325; Conservative 0; Mismatches 245; Indels 6; Gaps 2;
QY 58 CAAAGTCTTGGAGAAATGCGTGTGGCCACCACCGGAGCCGTCGCCGAGTTTACAGT 117
DB 9 CACCGTGATTGAGCAATGTCAAGTTGGCCACCCTCCCGCGGCGACGAGGTGATAC 68
QY 118 CCACATGTCTGTTTTCGACATCGATGGTGTATCTGTGATGCAGACACCATCTGCATT 177
DB 69 CCTCTTACTATTATTATGACCATGTTGGTTAGGTTTTCGCCGATGAGCGGATATTATT 128
QY 178 CTACAGATTCGGCCATCCTTCTCCCAACTC---TAAATTTATCATTTTCATCCATTAATC 234
DB 129 TTACAGCTCTCCATTTTCAACCCGATTTTCGTTCAAAACATATTCTCTCTTAAAA 188
QY 235 GTCCCTTTCCCTTGTCTCTCAACACACTTTCTTCGTTAGCCGGGAATTTGATTTGGCCGT 294
DB 189 TTCACCTCTCCCTCCTCCTCAACACATATAGCCGCTTAGCCGGAACCGTTGCTTGTCCACT 248
QY 295 AGATTCTCCGATAGATGCGGAGTTGCGTTACAGAAAGGGGACCTCCCTTTCTTTAAC 354
DB 249 AGATCAAAACG---GATATCTGAGTTACGTTATGTGACAGGAGATTTCTGTGCTGTTC 305
QY 355 AATTGAGAAATCGAGCATGATGATTTTATATATCTCGCGGAGATCATCAGAGGATTTCTTA 414
DB 306 TTTTTCGAGACTGATATGAATTTCAATATCTCATTTGGTGACCATCCGCGTAGGCTAA 365
QY 415 TAAATTAACGATTTGATTCGCGAGTGCAGCGATGTAACCTCCGCGGACGAGT 474
DB 366 GGATTTTATCACTTTGTTCTTAAGTTAGGGAACCTTAAGGATGTCACCGGGGTCCAACT 425
QY 475 ATTACCACTTTTTCCTTACAGGTGACGGTGTCTCAACACCGGTATATGATGACG 534
DB 426 AGCCCGCTCTTACGCAATTCAGGTGACACTTTTTCGGAATCTTTGGTGTATCCATTGGTT 485
QY 535 CAATCTTCATCAAGTCTTCTGTGATGCCAGTCTCTTTCTGCATTTTAAATAATATGGGT 594
DB 486 CACTAACCATCATGTTGTTGCTGATGGAACCTACTATAGCATGTTTCATTAAAGCGGGC 545
QY 595 TTTGGTTGACAAATCCAAATGAGATTCATTAAAGTT 630
DB 595 TTTGGTTGACAAATCCAAATGAGATTCATTAAAGTT 630

```

Db 546 TCTACTCCACAAATTACGTGGACATGACCAATCTT 581
RESULT 2
Bg599447
LOCUS EST504342 cSTS Solanum tuberosum cDNA clone cSTS5122 5' sequence, 12-APR-2001
DEFINITION mRNA sequence.
ACCESSION Bg599447
VERSION Bg599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 690)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
FEATURES
Location/Qualifiers
source 1..690
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS5122"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN

Query Match 9.3%; Score 151.4; DB 155; Length 690;
Best Local Similarity 57.1%; Pred. No. 1.9e-27;
Matches 316; Conservative 0; Mismatches 231; Indels 6; Gaps 2;

Qy 70 GAAATGCGTGTTCGGCCACCACCGGACGCGTTCGGAGTTTACAGTCCCACTGTCGTT 129
Db 2 GCAATGCAAGTTGGCCACCTCCGGGGCGGAGGAGGTGACACTCCCTCTACTTA 61
Qy 130 TTTGACATGCGATGGTGTCTCTGATGACAGAACACATCGATTTCTACAGATCCG 189
Db 62 TTTGATCATGTTTGGTTAGGTCTCCACCGPATACGGCGGATATTATTCTACAAGTCCC 121
Qy 190 CCATCCCTGTCCCACTC--TAATATTATCATTTTCATCCATTAAATCGTCCCTTCCCT 246
Db 122 CATTCCAAACCGGATTCGTTTCAAAACATTTATTCCTCCCTCTTAAATTCACCTCCCT 181
Qy 247 TGTCTCAACACACTTCTTCGTTAGCGGGAATTGATTGGCGGTTAGATTCTCCCGA 306
Db 182 CACTCTCAACACACTATAGCCCTTAGTGGAAACGTTGCTGCTCCACTAGATAAAGC- 240
Qy 307 TAGAATCGCGAGTTGGTTACAAAGAAGGGAGCTCCGTTTCTTTTAAACAATTCGAATC 366
Db 241 --GATATCCTGAGTTACATTATGTGACAGGAGATTCTGTATCTTACTTTTCTGAGAC 298
Qy 367 GAGCATGATTTTCAATATCTCCGCGGAGATCATCAGAGGATTTCTTATAAATTCACGA 426
Db 299 TGATATGATTTCAATTATCTCATTTGTTGACCATCCCGGTAATGCTAAGGATTTTATCA 358

```

```

Qy 427 TTTGATTCGCCAGCTGCCAGAACCGAATTGTAACTCTCGGCGACGAAGTATTACCACITTT 486
Db 359 CTTTGTTCCTACGTAGGGGAACCTAAGGATGCACCCGGGGTCCAATTAAGTCCGGTCTT 418
Qy 487 TCGTTTACAGGTGACGGTGTCTTCCCAACACCGGTATATGCAATTTGGACGCAATCTTCATCA 546
Db 419 AGCCATTCAAGTGACACTTTTTCGGAATCTTGGTGTATCCATTTGTTTCACTAACCATCA 478
Qy 547 AGTCTTGGTGTATGCCAGTTCTTTTCTGCATTTTAATAATATATGGGTTTGGTTGACAA 606
Db 479 TGTGTTGGTGTATGAGCTACCATAGTAGGTTCTATTAGGCGGTGGGCTCTACTCCATAA 538
Qy 607 ATCCAATGGAGAT 619
Db 539 ATTCGGTGGAGAT 551

RESULT 3
AW221049
LOCUS EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF3F1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 606)
AUTHORS Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,I., Craven,M.B., Bowman,C.L., Alm,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
Location/Qualifiers
source 1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F1"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN

Query Match 9.3%; Score 150.6; DB 113; Length 606;
Best Local Similarity 56.6%; Pred. No. 2.9e-27;
Matches 320; Conservative 0; Mismatches 239; Indels 6; Gaps 2;

Qy 58 CAAAGTTCTTGAGAAATGCCGTGTTGGCCACACCGGACCGCGTCCCGAGTTTACAGT 117
Db 9 CACCGTGATTGAGCAATGTCAAGTTGGCCACCTCCCGCGCGCGCAACGAGGTGATACT 68

```


FEATURES	Seq primer: M13F-R.	Location/Qualifiers
source	1. 763	
	/organism="Solanum tuberosum"	
	/cultivar="Kennebec"	
	/db_xref="taxon:4113"	
	/clone="BPL110A14"	
	/clone_lib="P. infestans-challenged leaf"	
	/tissue_type="leaf"	
	/dev_stage="6 week old"	
	/lab_host="SOLR"	
	/note="Vector: pBluescript SK(-): Site.1: EcoRI: Site.2: XhoI: Whole plants were challenged with 450,000 sporangia/ml P. infestans US-1(US 940501) in Blottron (Madison, Wisconsin). Leaf tissue was collected at 1, 2, 5, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR. Katahdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."	
BASE COUNT	212 a 162 c 156 g 233 t	
ORIGIN		
Query Match	8.7%;	Score 140.6; DB 155; Length 763;
Best Local Similarity	56.3%;	Pred. No. 9.8e-25;
Matches 346;	Conservative 0;	Mismatches 259; Indels 10; Gaps 4;
QY	75	GCCGTGTTGGCCACCACCGACCGCTC-GCCGAGTTTACAGTCCCACTGTCGTTTTC 133
Db	2	GTCAAGTTGGCCACCTCCGGCGCGCCAGCGAGGTGACACTCCCTCTACTATTATT 61
QY	134	GACATGCGATGGTGTCTCTGATGCGAGACACCATCTGCATTTCTACAGATTCGCGCAT 193
Db	62	GATCATGTTTGGTTAGGGTTCCACCGTATACGGCGGATATTATTCTACAAGTCGCCATT 121
QY	194	CCTGTGCCAA---CTCTAAATTTATCATTTCACTTAATCGTCCTTTCCCTTCTGTT 250
Db	122	TCCAACCCGATTCGTTCAAAACATTTATTCCTACTCTTAAAAATTCACCTCCCTCACT 181
QY	251	CTCAACACATTTCTTCGGTTAGCGGGAATTTGATTTGGCGGTAGATCTCCGATAGA 310
Db	182	CTCAACACTACACGCCCTTTAGCAGGAAATGTTCTGCTCCACTAGAT---ACAAATGGA 238
QY	311	ATGCCGAGTTGCTTTACAAGAAAGGGACTCCGTTCTTTTAACAATTTGCAGATCGAGC 370
Db	239	TATCCTGAGTTACGTTATGTGACAGGAAATCTGTACTGTACTTTTTTGACATGAT 298
QY	371	ATGGATTTTGATTATCTCGCCGGAGATCATCAGAGGATCTTTAAATTCACGATTTG 430
Db	299	ATGGATTTTCAATTATCTCATTTGGTGACCATCCGCGTAATGCTAAGGATTTTATCACATTT 358
QY	431	ATTCCGAGCTGCCAGAACCGATTGTAACTCCGGCGCAGGAAATATTACCACATTTTGGCT 490
Db	359	GTTCTTAAGTTAGGGGAACCTAAGGATGCACCCGGGTCCAAATTAGCTCCGGTCTTAGCC 418
QY	491	TTACAGTGTACGGTGTCTCCAACCCGGTATATGCAATTTGGACCGCAATCTTCATCAAGTT 550
Db	419	ATTCAAGTGACACTTTTCCGAATCTTGGTGATCCCATTTGGTTTCACTAACCATCATGTT 478
QY	551	CTTGGTGATGCCAGTCTTTTCTGCATTTTAAATAATATGGGTTTGGTTGACAAATCC 610
Db	479	GTTGGTGATGGAGCTACAGTAGTAGGGTTTCATTAAAGCGTGGGCTCTACTCCATAAATTC 538
QY	611	AATGGCATTCATTAAGTTTCCCTTCCACTTTCTCTACCTATGTAGCAGACATCTGTG 670
Db	539	GGTGGAGAT---GNAACAATCTTATCGAATGAGCTAATTCATTTTATGATAGTCCGTA 595
QY	671	GTCAAGATCCATTTT 685
Db	596	GTAAGGACCCATAT 610


```
QY 508 CTCAACACCCGGTATATGCAATTGGACGCAATCTTTCATCAAGATTCTTGGTGATGCCAGTTC 567
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 354 TCCGAATCTTGGTGATCAATGTTTTCACTAACCATCATCTGTTGGTGATGGAGTAC 413
QY 568 TTTTCTGCATTTTAATAAATATATCGGTTTGGTTGGTACAAATCCAAATCGAGATTCATTAAA 627
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 414 TATAGCAGGGTTCAATTAAGCGGTGGGCTCTACTCCACAATTCGGTGGACAT---GAACA 470
QY 628 GTTCTCCCACTTCTCTCTACCTATGTACGACAGATCTCTGGTGGGAAGATCCATTTC 687
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 471 ATTCTTATCGAATGAGCTAATTCATTTATGATAGGTCCTAGTAAAGAACCCCATATGG 530
QY 688 TATTTCGTGAAATAATCAATAATGAAGAAA 718
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 531 ACAAGGATGTCCTCTGGGAAGAAGTAA 561

RESULT 7
AW616206 AW616206 487 bp mRNA EST 24-MAR-2000
LOCUS hirsutum trichome, Cornell University Lycopersicon
DEFINITION EST307245 L. hirsutum trichome, Cornell University Lycopersicon
ACCESSION AW616206
VERSION AW616206.1 GI:7322240
KEYWORDS EST.
SOURCE Lycopersicon hirsutum.
ORGANISM Lycopersicon hirsutum.
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
REFERENCE 1 (bases 1 to 487)
AUTHORS van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from wild tomato (Lycopersicon hirsutum)
trichomes
TITLE Unpublished (2000)
JOURNAL Contact: David Frisch
COMMENT Lycopersicon hirsutum
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
    source
        1. 487
            /organism="Lycopersicon hirsutum"
            /db_xref="taxon:62890"
            /clone="cLHTD15"
            /clone_lib="L. hirsutum trichome, Cornell University"
            /tissue_type="trichome"
            /dev_stage="mixed stages"
            /note="Leaves of various stages were shaken in liquid
nitrogen, shearing off trichomes. This procedure yielded a
mixture of cells which is highly enriched for trichome
likely with minor contaminations of other types of leaf
cells"
BASE COUNT 116 a 117 c 91 g 163 t
ORIGIN

Query Match 7.98; Score 128.6; DB 118; Length 487;
Best Local Similarity 57.2%; Pred. No. 9.4e-22;
Matches 274; Conservative 0; Mismatches 199; Indels 6; Gaps 2;

QY 85 GCCACACCGGACCGTCGCGAGTTTACAGTCCCACTGCTGTTTTCGAGATCGCGATG 144
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GCCACCTCCAGCGCGGACGAGGTGGTACTCTCTCTTACTTATTTTGACCATGTTTG 60
QY 145 GTTGATCTCTGATGACAGAACACCATCTGCAATTTCTACAGATTCGCGCCATCCTTGTC 204
```

```
Db 61 GTTAGGGTTTCACCTTATTAGCGGATATATTCTACAAGCTCTCCATTTCCAAATCCGA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 205 --CTCTAAATTTATCATTTTCATCCATTAATCGTCCTTTCCCTTTGTTCTCAACACTT 261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 TTTCTGTTCAAAACATTAATTCCTCTCTTAAAAAATTCATTTTCCCTCACTCTCAAAACACTA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 262 TCTTCCGTTAGCCGGGAATTTGATTGGCCGGTAGATTTCCTCCGATAGAATGCGGGAGTT 321
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TATGCCCTTAGCCGGAAGGTTGCTTGTCACATAGATACAAACG--GATATCCTGAGTT 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 322 CGTTTACAAGAAAGGGGACTCCGTTTCTTTTAAACAATTGCAAGAATCGAGCAGGATTTTGA 381
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 238 ACGTTATGTGACAGAGATTCTGTCTGTTTAAATTTTCTGAGACTGATATGATGATTTCAA 297
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 382 TTATCTCGCGGAGATCATCAGAGGGATTCTTATAAATTCACGATTTCGCGCACT 441
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 298 TTATCTCATTTGGCGACCATCCCGTAATGCTAAGGATTTTATCACATTTGTTCTTAAGTT 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 442 GCCAAGACCGATTGTAACCTCCGGCGAGCAAGTATTACCACTTTTGTCTTTACAGGTGAC 501
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 358 AGGGAACCTAAGGATGCACCCGGGTCCAAATTAGCCCGCTCTTAGCCATTCAGGTGAC 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 502 GGTGTTCTCAACACCGGTATATGCAATTTGGACGCAATCTTCATCAAGTTCTTGGTGATG 560
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 418 ACTTTTCCGAATCTTGGTGATTCACATTTGGTTTCACTAACCATCATGTTGTTGGTGATG 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
AW930934 AW930934 497 bp mRNA EST 30-MAY-2000
LOCUS EST356777 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF42M23 5', mRNA sequence.
ACCESSION AW930934
VERSION AW930934.1 GI:8106335
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 497)
AUTHORS Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
    source
        1. 497
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone="cLEF42M23"
            /clone_lib="tomato fruit mature green, TAMU"
            /tissue_type="fruit pericarp"
            /dev_stage="mature green (3-5 days pre-ripening)"
            /lab_host="SOLR"
            /note="vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 132 a 106 c 109 g 150 t
```



```

Db 62 ACTTATTTTGATCATATATGTTGCGGTTTGGCTATATGCGTCGATTTTATTCTACAAG 121
Qy 185 TTCGCCACTCTTGCCCAA---CTCTAAATTTATCATTTTTCATCCATTAATATCGTCCCTT 241
Db 122 CTACCAATTTCCAAACTCGATTCGTTTCCAAACCATTAATCTCTACTCTTAAACATTCACCTC 181
Qy 242 TCCCTTGTTCTCAACACACTCTCTTCCGTTAGCCGGGAATTTGGATTTGGCCGGTAGATGCC 301
Db 182 TCCCTCACTCTCAACACACTACACACCTTAGCTGGCAACATTTGCTTGCCACTAAAT--- 238
Qy 302 TCCGATAGAAATGCGGAGTTCGCTTTACAGAAAGGGAGTCCGTTTCTTTTAAACAATTCGA 361
Db 239 TCGAGTGGTTATCTCTGAGTTCGCTTATGTGACTGGAGATTCATATATCTGTTACTTTATT 298
Qy 362 GAATCGACATGGATTTTGATTAATCTCCCGGAGATCATGAGGGGATCTTATATAATTC 421
Db 299 GAGACTGATATGAATTTCAATCACTCTCATTTGTTAACCATCTCGAAATGCTTAAGGATTTT 358
Qy 422 AACGATTTGATTCGCGAGTCCGAGAACCGATTTGTAACCTCCGGCGAGCAAGTATTACCA 481
Db 359 TATCCCTTCACTCTCAATTTGGCAACACTAAGGATGACCCGGGGTCAAACTAGTCCCG 418
Qy 482 CTTTGTGTTTACAGGTGACGGTGTCTTCCAAACCGGTATATGATGAGGACGCAATCTT 541
Db 419 GTCTTAGCCATTCAGTGACACTTTTTCGGAATCTTGGCATATCCGTTGGTTTAGTAAC 478
Qy 542 CATCAAGTCTTGTGTGAGCCAGTTC 567
Db 479 CATCATGTCGTTGTGATGGAATAC 504

RESULT 14
BE922746 574 bp mRNA EST 02-OCT-2000
LOCUS EST426515 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB21L12 5' sequence, mRNA sequence.
ACCESSION BE922746
VERSION BE922746.1 GI:10448822
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 574)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
Location/Qualifiers
1. 574
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB21L12"
/tissue_type="potato leaves and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
162 a 123 c 106 g 183 t

BASE COUNT
ORIGIN

```

```

Query Match 6.7% Score 108.8; DB 142; Length 574;
Best Local Similarity 54.9%; Pred. No. 9e-17;
Matches 281; Conservative 0; Mismatches 222; Indels 9; Gaps 3;

Qy 208 TAAATTTATCATATTCATCCATTAATATCGTCCCTTTCCCTTCTCTCAACACACTTTCTTCC 267
Db 63 TCAAACCATTTATTCCTACTCTTAACACATTCATCTCCCTCACTCTCAACACACTACACACC 122
Qy 268 GTTAGCCGGGAATTTGATTTGCGCGTAGATTTCTCCCATAGAAATGCCGGAGTTGCGGTTA 327
Db 123 CTTAGCTGGCAACATGCTTCTCCACTAAATTCGAGTGGTTAT---CCTGAGTTGCGGTTA 179
Qy 328 CAAGAAGGGGACCTCCGTTTCTTTAAACAATTCGAGAATTCGAGCATGGATTTTGATTAATCT 387
Db 180 TGTGACTGGAGATTTCTATATCTGTTACTTTTATTGAGACTGATATGAATTTCAATCACTCT 239
Qy 388 CGCGGAGATCATCAGAGGGGATTTCTTATAAATTCACCAATTTGATTCGCGAGCTGCCAGA 447
Db 240 CATTTGGTAACCATCTCTCGAAATGCTAAGGATTT---TTATCCCATTTCTCAATTTGGCACA 296
Qy 448 ACCGATTTAACCCTCCGCGGAGAGTATTACCACCTTTTTCGTTTACAGGTGACGGTGT 507
Db 297 ACCTAAGGATGACCGGGGGTCAAACTAGTCCGGTCTTAGCCCATTTCAAGTGACACTTTT 356
Qy 508 CTCCAACACCCGTTATATGCATTTGGACGCAATCTTCACTCAAGTTCTTGTGTGATGCCAGTTC 567
Db 357 TCCGATCTTGGCATATCCGTTGGTTTATTAGTAACCATGCTCGCTTGTGATGGAATAC 416
Qy 568 TTTTCTGCATTTTAATAAATTTGCGTTTGGTTTGACAAAATCCAATCCAGGATTCATTAAA 627
Db 417 CATCGTGAATTTCAAGAACAATGCGGTTTACTCAACAATTTCCGCGGTGAT---GAACA 473
Qy 628 GTTCTTCCACTTTCTCTCTACCTATGTACGACAGATCTCTGTGGTGAAGATTCATTCA 687
Db 474 GTGCTTAGAGAAATGAGTTTCATTTCCATTTTATGATAGGTCCGTGATAAAGACCCCTTATGA 533
Qy 688 TATTCGTCGAAAAATCTACAATGAAAGAAAAC 719
Db 534 ACAAGGGAGGATTTATATGGGATGAAATGAAGC 565

RESULT 15
BF053202 403 bp mRNA EST 16-OCT-2000
LOCUS EST438432 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB34D1 5' sequence, mRNA sequence.
ACCESSION BF053202
VERSION BF053202.1 GI:10807098
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 403)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
Location/Qualifiers
1. 403
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB34D1"
/tissue_type="potato leaves and petioles"

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source
1. 403
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB34D1"
/tissue_type="potato leaves and petioles"

```

/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 110 a 81 c 78 g 134 t
ORIGIN

Query Match 6.4%; Score 103.6; DB 143; Length 403;
Best Local Similarity 57.6%; Pred. No. 1.7e-15;
Matches 228; Conservative 0; Mismatches 159; Indels 9; Gaps 2;

QY	209	AAATTTATCATTTTCATCCATTAATCGCTTCCTTTCCTTCTCTCAACACATTTCTTC	CGG 268
Db	11	AAACCATTAATTCCTACTCTTAACATTCACCTCTCCCTCGCTCTCCAACACTATATG	ATT 70
QY	269	TTAGCCGGGAATTTGATTTGGCCGTAGATTCTCGATAGAAATGCGGAGTTGCGTTAC	328
Db	71	TTAGCCGGTAATGTGTTTATCGTTAAATTCAAACAACCTTA---CCTGAATTCGGTAT	127
QY	329	AAGAAAGGGGACTCCGTTTCTTTAACAATTCGAGAATCGAGCATGGATTTTGATTATC	TC 388
Db	128	GTAACAGGGGATTCGTGTCAGTTACTTTTCTGAAAGTGATGGATTTCAATTTTCTC	187
QY	389	GCCGGAGATCATCAGAGGGATTCTTATAATTCACGATTTGATTCGGCAGCTGCCAGAA	448
Db	188	ATAGGAGACCATCCTAGAAACGCAAGGATTTTATACATTTGTTCCAAAGTTGGCGGAA	247
QY	449	CC-----GATTGTAACCTCCGGCGACGAAGTATTACCACTTTTGTCTTACAGGTGACG	502
Db	248	CCTTGTAAGGCGACGGGGGATGTCCAATTAGCCCCCTTGTAGCCATTCAAGTGACA	307
QY	503	GTGTTCTCCAACACCGGTATATGATTTGGACGCAATCTTCATCAAGTCTGTGTGATGCC	562
Db	308	CTTTTCCGAATCATGCGCATATCGATTGGTTTCACCTAACCATCATGTTGTTGGTACGGA	367
QY	563	AGTCTCTTTCTGCATTTTATAAATTAATATGGGTTTG	598
Db	368	GCTAGCATAGGGTTTATAAAGCATGGGCTATG	403

Search completed: November 5, 2001, 18:00:43
Job time: 14883 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:03:41 ; Search time 168.74 seconds
(without alignments)
1819.738 Million cell updates/sec

Title: US-08-894-356C-2
Perfect score: 1622
Sequence: 1 GAACCATGTAATCAATTAA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2.6/ptodata/2/lna/5A_COMB.seq.*
2: /cgn2.6/ptodata/2/lna/5B_COMB.seq.*
3: /cgn2.6/ptodata/2/lna/6A_COMB.seq.*
4: /cgn2.6/ptodata/2/lna/6B_COMB.seq.*
5: /cgn2.6/ptodata/2/lna/PCTUS_COMB.seq.*
6: /cgn2.6/ptodata/2/lna/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	40.8	2.5	7218	1	US-08-232-463-14
2	39.6	2.4	51952	3	US-08-947-823-1
3	38	2.3	1953	1	US-08-463-975-1
4	38	2.3	1953	1	US-08-783-266-1
5	38	2.3	1953	2	US-09-023-327-1
6	37.6	2.3	19124	2	US-08-487-826B-13
7	37.6	2.3	636	4	US-08-998-416-1137
8	37.6	2.3	837	4	US-08-998-416-288
9	37.4	2.3	615	4	US-08-998-416-186
10	37	2.3	575	1	US-08-554-659-11
11	37	2.3	2821	4	US-09-103-429A-2
12	36.4	2.2	4160	4	US-09-341-678-5
13	35.8	2.2	894	3	US-08-906-769-90
14	35.8	2.2	894	3	US-08-906-616-90
15	35.8	2.2	894	3	US-08-817-795-90
16	35.8	2.2	894	3	US-08-639-075A-90
17	35.8	2.2	894	4	US-09-012-431-90
18	35.8	2.2	894	4	US-09-012-692-90
19	35.8	2.2	894	4	US-08-906-613-90
20	35.8	2.2	894	5	PCT-US95-14442A-90
21	35.6	2.2	602	1	US-08-764-100-8
22	35.6	2.2	642	1	US-08-764-100-13
23	35.6	2.2	643	1	US-08-764-100-7
24	35.6	2.2	2993	1	US-08-764-100-2
25	35.6	2.2	2993	1	US-08-764-100-10
26	35.6	2.2	3000	1	US-08-764-100-9
27	35.6	2.2	3001	1	US-08-764-100-1

28 35.6 2.2 152331 3 US-09-128-155-16 Sequence 16, Appl
29 35.4 2.2 720 1 US-08-117-083-23 Sequence 23, Appl
30 35.4 2.2 1121 1 US-08-207-412B-1 Sequence 1, Appl
31 35.4 2.2 1121 5 PCT-US95-02950-1 Sequence 1, Appl
32 35.4 2.2 1500 1 US-08-117-083-67 Sequence 67, Appl
33 35.2 2.2 961 6 5194596-16 Patent No. 5194596
34 35.2 2.2 961 6 5219739-16 Patent No. 5219739
35 35.2 2.2 972 3 US-09-286-690-1 Sequence 1, Appl
36 35 2.2 660 1 US-07-991-867B-32 Sequence 32, Appl
37 35 2.2 660 1 US-08-107-755A-32 Sequence 32, Appl
38 35 2.2 660 2 US-08-544-332-32 Sequence 32, Appl
39 35 2.2 1511 1 US-07-991-867B-8 Sequence 8, Appl
40 35 2.2 1511 1 US-08-107-755A-8 Sequence 8, Appl
41 35 2.2 1511 2 US-08-544-332-8 Sequence 8, Appl
42 35 2.2 2381 1 US-08-021-608D-9 Sequence 9, Appl
43 35 2.2 2381 1 US-08-726-160-9 Sequence 9, Appl
44 35 2.2 2381 5 PCT-US94-01782-9 Sequence 9, Appl
45 35 2.2 2384 1 US-08-021-608D-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935.313
; FILING DATE:
; APPLICATION NUMBER: EF 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZgpt-F1s
US-08-232-463-14


```
Query Match      2.3%   Score 38; DB 1; Length 1953;
Best Local Similarity 59.1%; Pred. No. 0.29;
Matches 65; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
```

QY 1513 TTATGATGCCAACAAATGTTTTAGAGAGATTACAAATCCATGGGAAGATCTATCAAACTCAT 1572
 ||||| ||||| ||||| ||||| ||||| |||||

Db 1844 TTTTAATGCRAATTTAATTTTGGAAATAAAAAGTCCCACGGTGTTTTGTTACTGTTTT 1903
 ||||| ||||| ||||| ||||| ||||| |||||

QY 1573 CTCCTCTATATATATATTCATTTGTTTTAAAAAASAAAAAAAAAAAAAA 1622
 ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Db 1904 CTCCATGAATAAACTCACTTGATTTTAAAAAASAAAAAAAAAAAAAA 1953
 ||||| ||||| ||||| ||||| |||||

RESULT 5

US-09-023-327-1

; Sequence 1, Application US/09023327

; Patent No. 5929225

; GENERAL INFORMATION:

; APPLICANT: WEI, ET AL.

; TITLE OF INVENTION: HAHB

; NUMBER OF SEQUENCES: 6

; CORRESPONDENCE ADDRESSES:

; ADDRESSEE: CARELLIA, BYRNE, BAIN, GILFILLAN,

; ADDRESSEE: CECCHI, STEWART & OLSTEIN

```
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

Query Match      2.3%; Score 38; DB 2; Length 19124;
Best Local Similarity 53.3%; Pred. No. 0.97;
Matches 80; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 1473 AAGAAAAGTGGTATCAATGTATATAAAAGACAGACAAAGTTATGATGCAACAAATGTTT 1532
Db 15686 AAAATATATATCATATAAACAACAAAAAAGAAAAAATATATTAATAAATAATATA 15745
QY 1533 TAGGAGATTACAAATCCATGGAAGATGTATCAAACTCATCTCTATATATATATATTC 1592
Db 15746 TCATAAATAAATAAATAATTAATAAATAATTAATAAATAAATAAATAAATAA 15805
QY 1593 AATTGTTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1622
Db 15806 AAATTTTAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 15835

RESULT 7
US-08-998-416-1137/c
; Sequence 1137, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38,241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CGC1976
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1692RP
; US-08-998-416-1137

Query Match      2.3%; Score 37.6; DB 4; Length 636;
Best Local Similarity 55.3%; Pred. No. 0.21;
Matches 73; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 1490 AATGTATATAAAAGACAGACAAAGTTATGATGCAACAAATGTTTATGAGGATTTACAAATCC 1549
Db 443 ATTAATAATATAAATCAACATAATATTTATAAATAAGATATATATAAATAAATAAATAA 384
QY 1550 ATGGGAAGATGTATCAAACTCATCTCTATATATATATATATATATATATATATATATAT 1609
Db 383 TTTACAATATTTAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 324
QY 1610 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1621
Db 323 ATTAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 312

RESULT 8
US-08-998-416-288/c
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Rebschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
```

```

; REFERENCE/DOCKET NUMBER: PF/5-5-30306/A/CGC1976
;
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587
; TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 186:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: PAG1074RP
; US-08-998-416-186

```

Query Match	2.3%	Score 37.4:	DB 4:	Length 615:
Best Local Similarity	55.5%	Pred. No. 0.23:		
Matches	71:	Conservative	0:	Mismatches 56: Indels 0: Gaps 0:
Qy	1495	ATAAAAAGACAGACAAGTTATGATGCAACAACAACTCTTTTAGCAGATTACAAATCCATGGG	1554	
Db	438	ATAATATAAATCAACATAAATATTTATAAAATAGATATTATAATAAAATAATAATTTAC	379	
Qy	1355	AAGATGTATCAAACTCATCTCTATATATATATATTCAAATTTGTTTTAAAAAIAAAAAAAAA	1614	
Db	378	ANTATTTAAATAAATAATCTTTTATAATAATAAAATATATATTTTTTAATAACAATTAA	319	
Qy	1615	AAAAAA	1621	
Db	318	TAAATAA	312	

```

RESULT 10
US-08-554-659-11
: Sequence 11, Application US/08554659
: Patent No. 5767361
: GENERAL INFORMATION:
: APPLICANT: Wingate, Vincent
: APPLICANT: Wolff, Mark
: TITLE OF INVENTION: LEPIDOPTERAN GABA GATED CHLORIDE CHANNEL
: TITLE OF INVENTION: AND METHODS OF USE THEREOF
: NUMBER OF SEQUENCES: 12
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Scully, Scott, Murphy & Presser
: STREET: 400 Garden City Plaza
: CITY: Garden City
: STATE: New York
: COUNTRY: USA
: ZIP: 11530-0299
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/554,659
: FILING DATE:
: CLASSIFICATION: 436
: ATTORNEY/AGENT INFORMATION:
: NAME: Digiglio, Frank S
: REGISTRATION NUMBER: 31,346
: REFERENCE/DOCKET NUMBER: 9732
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (516) 742-4343
: TELEFAX: (516) 742-4366
: TELEX: 230 901 SANS UR
: INFORMATION FOR SEQ ID NO: 11:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 575 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: single
: TOPOLOGY: linear

```

US-09-103429A-2
TISSUE TYPE: peritrophic membrane


```

; APPLICANT: Rushlow, Keith E.
; APPLICANT: Hunter, Shirley Wu
; APPLICANT: Frank, Glenn R.
; APPLICANT: Heath, Andrew W.
; APPLICANT: Yamaka, Miles Yamanaka
; APPLICANT: Arfsten, Ann
; APPLICANT: Dale, Beverly
; APPLICANT: Stiegler, Gary
; TITLE OF INVENTION: USE OF PROTEASE INHIBITORS AND
; TITLE OF INVENTION: PROTEASE VACCINES TO PROTECT ANIMALS FROM FLEA
; TITLE OF INVENTION: INFESTATION, AND FLEA PROTEASE PROTEINS, NUCLEIC ACID
; TITLE OF INVENTION: MOLECULES, AND USES THEREOF
; NUMBER OF SEQUENCES: 119
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Sheridan Ross & McIntosh
; STREET: 1700 Lincoln Street, Suite 3500
; CITY: Denver
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80203
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/817,795
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/14442
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Gary J. Connell
; REGISTRATION NUMBER: 32,020
; REFERENCE/DOCKET NUMBER:
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 90:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 894 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..766
; FEATURE:
; NAME/KEY: Xaa - any amino acid
; LOCATION: 187
; US-08-817-795-90

Query Match 2.2%; Score 35.8; DB 3; Length 894;
Best Local Similarity 49.7%; Pred. No. 0.81;
Matches 91; Conservative 0; Mismatches 92; Indels 0; Gaps 0;

QY 1439 TTCATCAATAGTCTCCATGTGAGGAGCCCTTTCTAGAAAAAAGTGGTATCAATGTATAA 1498
   ||||| ||| | ||||| ||| | ||||| ||| | ||||| ||| | ||||| ||| |
Db 709 TTACCAATGTTGCTCATTTTGTGGATTGGATTCAAGAACATAGTGGATTGGAATTGTA 768
   ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
QY 1499 AAAAGACAGACAAGTATGATGCAACAAATGTTTATAGGAGATTACAAATCCATGGGA 1558
   ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 769 AATAAACTCAAACTAAAAAATAATTTAATGCACTGAAAAATTTTCAAGAAAA 828
   ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
QY 1559 TGTATCAAACTCATCTCTATATATATATATCAATTTTAAAAAATAAAAAA 1618
   ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
Db 829 GTTTGGATCGTTTGTAAATGTAATGACATAAAGCGTAAATTAGAAAAA 888
   ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| | ||| |
QY 1619 AAA 1621
   |||

```

Db 889 AAA 891

Search completed: November 5, 2001, 18:04:15
Job time: 15055 sec

THIS PAGE BLANK (uspto)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:15 ; Search time 445.49 seconds
(without alignments)
2286.149 Million cell updates/sec

Title: US-08-894-356c-2

Perfect score: 1622

Sequence: 1 GAACCAATGAATCAATTAA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*

1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1607.6	99.1	1622	17 AAT37309
2	218	13.4	1479	17 AAT37311
3	190.8	11.8	1703	17 AAT37308
4	153.8	9.5	1518	17 AAT37313
5	139.8	8.6	1508	17 AAT37312
6	96.8	6.0	936	22 AAF58252
7	96.8	6.0	936	22 AAF58254
8	96.8	6.0	936	22 AAF58257
9	96.8	6.0	936	22 AAF58259
10	96.8	6.0	936	22 AAF58262
11	96.8	6.0	938	22 AAF58255

12	95.4	5.9	936	22	AAF58252	Oligonucleotide D1
13	95.4	5.9	935	22	AAF58254	Oligonucleotide D1
14	95.4	5.9	935	22	AAF58257	Oligonucleotide D1
15	95.4	5.9	935	22	AAF58259	Oligonucleotide D2
16	95.4	5.9	936	22	AAF58262	Oligonucleotide D2
17	95.4	5.9	938	22	AAF58255	Oligonucleotide D1
18	58.4	3.6	244	22	AAF58238	Oligonucleotide D1
19	56.2	3.5	244	22	AAF58238	Oligonucleotide D1
20	47.8	2.9	2969	21	AA23450	cDNA encoding huma
21	47.2	2.9	1240	21	AAC53227	Arabidopsis thalia
22	45.8	2.8	328	21	AA331809	Plant microsatelli
23	43.8	2.7	280	16	AAT21000	Human gene signatu
24	43.8	2.7	1273	21	AA80571	Human secreted pro
25	43.4	2.7	1595	10	AA90733	Genomic DNA from P
26	43.4	2.7	5376	10	AA90732	Sequence of plasmi
27	42.4	2.6	7458	21	AA70106	Plasmodium falcipa
28	42	2.6	1982	10	AA90225	Malaria-specific P
29	41.4	2.6	471	21	AA16053	Human colon cancer
30	41.4	2.6	1523	21	AA59768	Human secreted pro
31	41.4	2.6	1823	21	AA78037	Human cancer assoc
32	41.4	2.6	2851	20	AA233599	Human breast tumou
33	41.2	2.5	3975	9	AA81157	Malaria-specific g
34	41.2	2.5	3975	13	AAQ22999	SERP gene. Plasm
35	41.2	2.5	6124	11	AAQ03568	Sequence encoding
36	40.8	2.5	9789	17	AAT41852	cDNA encoding Plas
37	40.4	2.5	769	18	AAT79095	Full length human
38	40	2.5	13499	21	AAF22286	BAC containing rep
39	39.8	2.5	498	22	AAF68733	Human lung tumour
40	39.8	2.5	969	21	AA77818	Human cancer assoc
41	39.8	2.5	3366	22	AAF24684	Nucleotide sequenc
42	39.8	2.5	3366	22	AAF24706	Nucleotide sequenc
43	39.8	2.5	3515	22	AAF75343	Human TGF-beta rec
44	39.8	2.5	10442	22	AAF24680	Nucleotide sequenc
45	39.8	2.5	10442	22	AAF24702	Nucleotide sequenc

ALIGNMENTS

RESULT 1

AAT37309
ID AAT37309 standard; cDNA to mRNA; 1622 BP.

XX AC AAT37309;

XX DT 06-FEB-1997 (first entry)

XX DE Aromatic acyl transferase coding sequence.

XX KW Aromatic acyl transferase; transformatio; anthocyanin pigment;

XX KW Plants; acylation; colour; tone; colouration; colour change;

XX KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;

XX KW Scenecio cruentus; Lavandula angustifolia; ds.

XX OS Gentiana triflora var. japonica (Clone pgAT106).

XX FH Key Location/Qualifiers

XX CDS 6..1415

XX FT /*tag= a

XX FT /product= Aromatic acyl transferase.

XX FT 701..703

XX FT /*tag= b

XX FT /transl_except= ATC encodes Leu

XX FT 1322..1324

XX FT /*tag= c

XX FT /transl_except= AAT encodes Asp

XX PN WO9625500-A1.

XX XX 22-AUG-1996.

XX PF 16-FEB-1996; 96WO-JP00348.

XX XX

Db 1027 gcyttgcaaaagcaacacataaaagagtagttggggataaaagggtcttctgttgcaagtg 1086
 Qy 1119 CGCCATCGGAGATGCTATCCATAGAGGTTACATGACTACGAGGAATTCGAGAGAG 1178
 Db 1087 cagctattggagaagccattgaaaagaggttgcaacaacgaaaagcgtctcttgagatg 1146
 Qy 1179 ATGTGATATCGCGCCCGGATCAACATCTGCGGCACCAAGGTCGACGCTCATTTATGTCG 1238
 Db 1147 caaaaacttggtatcggaatcaat-----ggaatccctcaaaaagattctcg 1197
 Qy 1239 TTGGATCGGCACACGCAATGTCAGATGATTTGATGTCAGATTTTGGTGGGAAAGCTTG 1298
 Db 1198 ggaattaccggatcgctcaagttcgatcgatggtgtagatttgatgggggaaagcctg 1257
 Qy 1299 AAACGATGATCTGTTCAACTATCTTCGGCAACACTAATTTTGATCTCTCGGTCCA 1358
 Db 1258 caaatttgaca---ttacctgttgattatgcagaattgattatgattcgatgcagtcga 1314
 Qy 1359 GAAGATTTAAAGGAGCAGCTTGAGCTTGGCATTTCTTTGCTTAAGATAGGATGGACGCAT 1418
 Db 1315 gggatttgaaaaaggttgagatggagatgcattcgttaagattcattatgagatgcatt 1374
 Qy 1419 TTGCCACCATTTT 1432
 Db 1375 ttgcaaaaatcttt 1388

RESULT 4

AAT37313
 ID AAT37313 standard; cDNA to mRNA; 1518 BP.
 XX
 AC AAT37313;
 XX
 DT 06-FEB-1997 (first entry)
 XX
 DE Aromatic acyl transferase coding sequence.
 KW Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla octinoides;
 KW Scenecio cruentus; Lavandula angustifolia; ds.
 XX
 OS Lavandula angustifolia (Clone pLAT21).
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1355
 FT /*tag= a
 FT /product= Aromatic acyl transferase.
 XX
 PN W09625500-A1.
 XX
 PD 22-AUG-1996.
 XX
 PF 16-FEB-1996; 96WO-JP00348.
 XX
 PR 30-JAN-1996; 96JP-0046534.
 PR 17-FEB-1996; 95JP-0067159.
 PR 29-JUN-1996; 95JP-0196915.
 XX
 PA (SUNR) SUNTORY LTD.
 XX
 PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX
 DR WPI; 1996-393401/39.
 XX
 PT DNA coding for aromatic acyl transferase - for transforming plants
 PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers
 XX
 PS Claim 4; Page 73-76; 94pp; Japanese.
 XX

CC Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313. NOTE: This
 CC sequence is supposed to cross reference with the protein described
 CC in AAW04727, however there are so many discrepancies between the
 CC polypeptide decoded from this sequence and the polypeptide given in
 CC the specification and described in AAW04727 that the indexer decided
 CC not to cross reference the two.
 XX

SQ Sequence 1518 BP; 384 A; 340 C; 366 G; 428 T; 0 other;

Query Match 9.5%; Score 153.8; DB 17; Length 1518;
 Best Local Similarity 49.5%; Pred. No. 8.4e-30;
 Matches 686; Conservative 0; Mismatches 642; Indels 57; Gaps 9;

Qy 72 AATGCGGTGTTGCGCCACCACCGGACGCGTCCGAGTTTACAGTCCACATGTCGTTTT 131
 Db 16 aatcccgagtgccgcccctccagcgagtggtgagcagtcactcccgccactctct 75
 Qy 132 TCGACATCGATGCTGATCTCTGATGCAGAACACATCTGCAATTTTACAGATTCGCCGC 191
 Db 76 tcgacatgacgtggtgcaattccaccccatgcttcagcttctcttcagaaacctct 135
 Qy 192 ATCTTGTCGCCAATCTA---AATTTATCATTTTCATCCATTAATCGTCCCTTTCCCTTG 248
 Db 136 gttccaaacccgcttctctgaaacgctgttccgaaacactcaacaactctctctctaa 195
 Qy 249 TTCTCAACACATTTCTCCGTTAGCGGGAATTTGATTTGGCGGTAGATTTCTCCGATA 308
 Db 196 cccctcaaacactctctccctcttcgatgcactcactcaccctctctctctctctct 252
 Qy 309 GAATGCGGAGTTGGCTTACAAAGAAAGGGGACTCCGTTTCTTTAAACAATTCAGCAATCA 368
 Db 253 aaatgccgagttccggtatcagaacggtgactggttctcttcacgattatggagctg 312
 Qy 369 GCATGGATTTTGAATATCTCGCGGAGATCATCAGAGGATTTCTTATAAATTCACAGATT 428
 Db 313 tcggagatcatccgcatctccataactactgctttgccccctagcgcgattatg 372
 Qy 429 TGATTTCCGCGAGCTGCAGAACCGATTGTAACCTCGCGGACGAAGTATTACACATTTTGG 488
 Db 373 aagatctccagctgcccgcgagtagtcgaggaatctgtagcggaaattgtttcagatttag 432
 Qy 489 CTTTACAGGTGACGGTGTCTCCAAACACCGGTATATGCAATTCGACGAATTCATCAAG 548
 Db 433 ccgtgcaagtactctgtttcccggtccggtggtgcatcggaataacgacgcaccaca 492
 Qy 549 TTCTTGGTGATGCCAGTTCTTTCTGCAATTTTAAATAATATATGTTGGTTTGGTTGACAAAT 608
 Db 493 ccgttagcagatgctccatcgttttagggtttatgaagagttgggtccctccactcaaat 552
 Qy 609 CCAATGGAGATTC---ATTAAAGTTCCCTCCACTTCTCTACCTATGATGACGACAGAT 665
 Db 553 tcggaggagatgatgaattcttggacgggaaaggtagtcttgcgggttttcgacgcat 612
 Qy 666 CTGTGTGTCGAAGATCCATTTTCATATTCGTGCAAAATCTACAAATGAAGAAACTCTCA 725
 Db 613 cgtcgtgtaattatccgcttaaatggacacataattatggaaac-----aacgcgcaga 666
 Qy 726 AATCTCAGGGCACCTACTGTTCTAAATCCAGCAATTTCTAAAGATGAAGTTTCGAGCCA 785
 Db 667 aacgtccgttggaaatc-----gcagcatccatctttaccgcgcatcggttcagacta 720
 Qy 786 CCTTCATCTACACCTATTGATATCATCAAGCTCAAGAAATTCATTTCTGTCAAAAATC 845
 Db 721 cctacctttccacccaatctgaaattgaaaggtttgattcagagaaagcc 780
 Qy 846 GCAACTTAACCGGTAGTAGTAATATAATCTGTCAACTTTTACGGGTGACATCTGCACCTGA 905

[illegible]

RESULT

AAE58252/C

ID AAF58252 standard; DNA; 936 BP.

XX

AC AAF58252;

XX

DT 24-APR-2001 (first en

XX

DE Oligonucleotide D1835.

XX

KW Electron-transfer grou

KW gene expression; ss.

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

OS Synthetic.

XX
150300107000-10

PN WO200107665-A2.

 $\gamma\gamma$

PD	01-FEB-2001.	
XX		
XX	26-JUL-2000; 2000WO-US20476.	
XX		
PR	26-JUL-1999; 99US-0145695.	
PR	17-MAR-2000; 2000US-0190259.	
XX		
XX	(CLIN-) CLINICAL MICRO SENSORS INC.	
XX		
XX	Umek RM;	
XX		
XX	WPI; 2001-159728/16.	
XX		
PT	Nucleic acids containing electron-transfer group, useful as labels in	
PT	hybridization assays, e.g. for genotyping, allowing repeat analyses on	
PT	a single surface	
XX		
PS	Example 6; Page 127; 159pp; English.	
XX		
XX	The present invention relates to a composition comprising two nucleic	
CC	acids each containing an electron-transfer group (ETM) having	
CC	different redox potentials. The invention is used for electronic	
CC	detection of nucleic acids, especially of substitutions (mismatches)	
CC	and single-nucleotide polymorphisms, e.g. for genotyping,	
CC	monitoring gene expression.	
XX		
SQ	Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;	

Query Match	6.0%	Score 96.8;	DB 22;	Length 936;
Best Local Similarity	1.0%;	pred. No. 2.8e-15;		
Matches	8;	Conservative 453;	Mismatches 305;	Indels 0; Gaps

[illegible]

	Qy	1457	GTGAGGACCCCTTTGTAAGAATAAACTGGTATCAATGTATATAAAAAGACAGACAAGTAT	15151
	Dd	190	WWWWWWW	131
	Qy	1517	GATGCACAAATCCTTTTAGGAGATTACAAATCCATGGGAAGATGTATCAAACATCATCTCT	1576
	Dd	130	WWWWWWW	71
	Qy	1577	CTATATATATATTCAAATGTTTTTAAAAAATAAAAAAAAAAAAAA	1622
	Dd	70	WWWWWWW	25
	RESULT 7			
	AAF58254/c			
	ID	AAF58254 standard; DNA; 936 BP.		
	XX AC	AAF58254;		
	XX DT	24-APR-2001 (first entry)		
	XX DE	Oligonucleotide D1875.		
	XX KW	Electron-transfer group; ETM; mismatch; genotyping;		
	XX KW	gene expression; ss.		
	XX OS	Synthetic.		
	XX PN	WO200107665-A2.		
	XX PD	01-FEB-2001.		
	XX PF	26-JUL-2000; 2000WO-US20476.		
	XX PR	26-JUL-1999; 99US-0145695.		
	XX PR	17-MAR-2000; 2000US-0190259.		
	XX PA	(CLIN-) CLINICAL MICRO SENSORS INC.		
	XX PI	Umek RM;		
	XX DR	WPI; 2001-159728/16.		
	XX PT	Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -		
	XX PS	Example 6; Page 127; 159pp; English.		
	XX CC	The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.		
	XX SQ	Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;		
	Query Match 6.0%; Score 96.8; DB 22; Length 936;			
	Best Local Similarity 1.0%; Pred. NO. 2.8e-15;			
	Matches 8; Conservative 453; Mismatches 305; Indels 0; Gaps			
	Qy	857 GG TAGTAGTAA TTATTA ATCGTCAACTTCACGGTGACATCTGCCACTGATCGACATGC	916	
	Dd	GGC	731	
	Qy	TTCGCAAAATCATTAGCACACCGCTCGTAAAGAGAGAGATAAACATGCACGA	976	
	Dd	WWWWWWW	671	
	Qy	AAC TTATGTCTTTCATCAACTGCCGACAAACGTTTGTCTCCCGGATACCCTCAAATTA	1036	

[illegible]

Synthetic.

WO200107665-A2.

01-FEB-2001.

26-JUL-2000; 2000WO-US20476.

26-JUL-1999; 99US-0145695.

17-MAR-2000; 2000US-0190259.

(CLIN-) CLINICAL MICRO SENSORS INC.

Umek RM;

WPI; 2001-159728/16.

Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -

Example 6; Page 127; 159pp; English.

The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.

Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;

Query Match 6.0%; Score 96.8; DB 22; Length 938;
Best Local Similarity 1.0%; Pred. NO. 2.8e-15;
Matches 8; Conservative 453; Mismatches 305; Indels 0; Gaps 0;

QY 857 GGTAAGTATGCTCAACTTCACGGTGACATCGCACATGCTGGACATGC 916
|| :
DB 790 GCWWWWWWWMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMMM 731
QY 917 TTGTGGAATCATTAGACACCCTCGTAAGACAGAAGTGAAGAGATAAACATGCAGCA 976
:
DB 730 WWWWWWMM 671
QY 977 AACATTATGCTTTCAATCAACATGCCGACAACGTTTTGCTCCGCGATACCTCAAATPAC 1036
:
DB 670 WWWWWWMM 611
QY 1037 TTTGGAATTCATAGTCCTTTGATGGTGGATCGACATCATGACCACTTGTAGGAAT 1096
: : : : : : : | : : : : : : : : : : : : : : : :
DB 610 WWWWWWMM 551
QY 1097 GAAGGGTTGTGCGTAGCTGCAACCGCCCATCGAGATGCTATCCATAAGAGTTACATGAC 1156
:
DB 550 WWWWWWMM 491
QY 1157 TAGCAAGAAATTCGAGAGAGATGGATATCGCGCCCCCAGATCAACATCTCGCGCACCA 1216
:
DB 490 WWWWWWMM 431
QY 1217 AGGTCGACGCTCATTTATGTCGTTGGATCCGACACGCAATGTGCATGATTTTCATGCA 1276
: : : : : : : : : : : : : : | : : : : : : : : : :
DB 430 WWWWWWMM 371
QY 1277 GATTTTGGTGGGAAAAGCTTGAAAGCATGAATCTGTTTCAACTAATCTTCGGGCAACA 1336
:
DB 370 WWWWWWMM 311
QY 1337 CTAAATTTGATCTCTCGGTCCAGAAGATTTAAAGAGGACACTTGACATGCGATTCCTTGT 1396
:
DB 310 WWWWWWMM 251

CC detection of nucleic acids, especially of substitutions (mismatches)
 CC and single-nucleotide polymorphisms, e.g. for genotyping,
 CC monitoring gene expression.

SQ Sequence 936 BP; 6 A; 138 C; 8 G; 8 T; 776 other;

```

Query Match          5.9%; Score 95.4; DB 22; Length 936;
Best Local Similarity 1.5%; Pred. No. 6.3e-15;
Matches 12; Conservative 451; Mismatches 312; Indels 0; Gaps 0;

Qy 128 TTTTTCGACATGCGAGTTCATCTCTGATGCGAAGACCATCTGCAATTTCTACAGATTC 187
Dy 24  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 24  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 188 CGCATCCTTGTCCCAACTCTAAATTTATCATTTTCATCCATTAATCGTCCCTTCCCTT 247
Dy 84  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 84  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 248 GTTCTCAACACATTTCTTCCGTAGCCGGGAATTTGATTTGGCCGGTAGATTCCTCCGAT 307
Dy 144 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 144 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 308 AGAATGCCGGAGTTCGCTTACAAAGGAGGACCTCCGTTCTTTAAACAATTCAGAAATCG 367
Dy 204 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 204 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 368 AGCATGGATTTGATTAATCTCGCCGGAGATCATCAGAGGGATTTCTTATAAATTCACGAT 427
Dy 264 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 264 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 428 TTGATTCGCGAGCTGCCAGAACCGATTTGAACCTCCGGCGAGAGTATTACCACTTTT 487
Dy 324 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 324 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 488 GCTTACAGGTGAGGGTTCCTCAACACCGGTATATGCATTTGGACGCAATCTTCATCAA 547
Dy 384 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 384 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 548 GTTCTTGGTGATGCCAGTTCCTTCTGCAATTTAATAAATTTATGGGTTTGGTTGACAAA 607
Dy 444 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 444 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 608 TCCATGGAGATTCATTAAGTTCCTTCCACTTTCTTCTTACCTATGTACGACAGATCT 667
Dy 504 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 504 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 668 GTGGTCAAGATCCATTTTCATATTCGTCGGAATAATCTACAATGAAGAAACTGCTCAA 727
Dy 564 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 564 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 728 TCTCAGGGCACCTACTGTCTTAAATCCAGCAATTTCTAAAGATGAAGTTCGAGCCACC 787
Dy 624 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 624 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 788 TTCATCTTACACCTATTTGATATCATGAGCTCAAGAAATTCATTTCTGTCAAAAAATCGC 847
Dy 684 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 684 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 848 AACTTAACCGGTAGTAGTAATTAATCTGTCAACTTTCACGGTGACATCTGCAC 902
Dy 744 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Dy 744 : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Search completed: November 5, 2001, 18:11:26
 Job time: 15421 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model
Run on: November 5, 2001, 16:30:46 ; Search time 8904.87 Seconds
(without alignments)
2817.412 Million cell updates/sec

Title: US-08-894-356c-2
Perfect score: 1622
Sequence: 1 GAACCATTTGAATCCAATTAA.....AAAAAAAAAAAAAAAAAAAA 1622

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl: *
1: gb_ba1: *
2: gb_ba2: *
3: gb_ba3: *
4: gb_in1: *
5: gb_in2: *
6: gb_in3: *
7: gb_om: *
8: gb_ov: *
9: gb_pat1: *
10: gb_pat2: *
11: gb_ph: *
12: gb_pl1: *
13: gb_pl2: *
14: gb_pl3: *
15: gb_pl4: *
16: em_ba1: *
17: em_ba2: *
18: em_fun: *
19: em_htgo_hum: *
20: em_htgo_inv: *
21: em_htgo_rod: *
22: em_htg_hum1: *
23: em_htg_hum2: *
24: em_htg_hum3: *
25: em_htg_hum4: *
26: em_htg_hum5: *
27: em_htg_hum6: *
28: em_htg_hum7: *
29: em_htg_hum8: *
30: em_htg_inv1: *
31: em_htg_inv2: *
32: em_htg_other: *
33: em_htg_rod: *
34: em_hum1: *
35: em_hum2: *
36: em_hum3: *
37: em_hum4: *
38: em_hum5: *
39: em_hum6: *
40: em_hum7: *
41: em_in: *
42: em_om: *
43: em_or: *

44: em_ov: *
45: em_pat: *
46: em_ph: *
47: em_pl: *
48: em_ro: *
49: em_sts: *
50: em_sy: *
51: em_un: *
52: em_vi: *
53: gb_sts1: *
54: gb_sts2: *
55: gb_sts3: *
56: gb_sy: *
57: gb_un: *
58: gb_v1: *
59: gb_v2: *
60: gb_htg1: *
61: gb_htg2: *
62: gb_htg3: *
63: gb_htg4: *
64: gb_htg5: *
65: gb_htg6: *
66: gb_htg7: *
67: gb_htg8: *
68: gb_htg9: *
69: gb_htg10: *
70: gb_htg11: *
71: gb_htg12: *
72: gb_htg13: *
73: gb_htg14: *
74: gb_htg15: *
75: gb_htg16: *
76: gb_htg17: *
77: gb_htg18: *
78: gb_htg19: *
79: gb_htg20: *
80: gb_htg21: *
81: gb_htg22: *
82: gb_htg23: *
83: gb_htg24: *
84: gb_htg25: *
85: gb_pr1: *
86: gb_pr2: *
87: gb_pr3: *
88: gb_pr4: *
89: gb_pr5: *
90: gb_pr6: *
91: gb_pr7: *
92: gb_pr8: *
93: gb_pr9: *
94: gb_rol: *
95: gb_rol2: *
96: gb_in4: *
97: gb_pr10: *
98: em_ba3: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DB ID	Description
1	1622	100.0	1622	10	E12754
2	1622	100.0	1622	12	AB026494 Gentiana
3	218	13.4	1479	10	E12756
4	216	13.3	1476	12	AB029340 Perilla f
5	189.2	11.7	1679	12	AB010708 Gentiana
6	189.2	11.7	1703	10	E12753
7	139.8	8.6	1508	10	E12757
8	50.4	3.1	178273	60	AC005308 Plasmodi

c 9	50.2	3.1	205429	60	AC005506	AC005506 Plasmid
c 10	50	3.1	169794	60	AC004688	AC004688 Plasmid
c 11	50	3.1	196149	60	AC004709	AC004709 Plasmid
c 12	49	3.0	42781	14	MISOLSUR	X17375 Scenedesmus
c 13	49	3.0	42919	13	AF204057	AF204057 Scenedesmus
c 14	49	3.0	234112	96	PFMAL4P2	AL035475 Plasmid
c 15	48.2	3.0	144457	67	AC023268	AC023268 Homo sapi
c 16	48	3.0	101176	12	AC002560	AC002560 Genomic s
c 17	46.8	2.9	1337	88	AF090922	AF090922 Homo sapi
c 18	46.8	2.9	66237	12	AB016892	AB016892 Arabidops
c 19	46.6	2.9	85779	14	SC5011856	AJ011856 Saccharom
c 20	46.6	2.9	178273	60	AC005308	AC005308 Plasmid
c 21	46.4	2.9	153477	60	AC006278	AC006278 Plasmid
c 22	46.4	2.9	192581	84	PFMAL13P1	AL049180 Plasmid
c 23	46.2	2.8	83110	84	PFMAL13PB	AL109814 Plasmid
c 24	46.2	2.8	293431	84	PFMAL13P4	AL049181 Plasmid
c 25	46	2.8	192676	61	AC010888	AC010888 Homo sapi
c 26	45.8	2.8	156550	60	AC015830	AC015830 Homo sapi
c 27	45.4	2.8	1063	53	CNS07A2Y	AL436064 T3 end of
c 28	45.2	2.8	2538	14	PSXYNAGEN	X91858 Ptiomyces s
c 29	45.2	2.8	53932	68	AC023371	AC023371 Homo sapi
c 30	45.2	2.8	78379	12	AB006696	AB006696 Arabidops
c 31	45.2	2.8	160877	70	AC026560	AC026560 Homo sapi
c 32	45.2	2.8	174303	74	AC069506	AC069506 Homo sapi
c 33	45	2.8	5560	96	PFAL132422	AJ132422 Plasmid
c 34	44.8	2.8	5917	96	PFANT195	X02919 Plasmid
c 35	44.8	2.8	164399	96	PFMAL3P6	Z98551 Plasmid
c 36	44.6	2.7	35928	6	CELFALH10	U61954 Caenorhabdi
c 37	44.6	2.7	119918	82	AP000713	AP000713 Homo sapi
c 38	44.6	2.7	173805	91	CNS05TBW	AL352977 Human chr
c 39	44.6	2.7	293431	84	PFMAL13P4	AL049181 Plasmid
c 40	44.4	2.7	897	53	CNS07ABZ	AL436389 T7 end of
c 41	44.4	2.7	1032	53	CNS06LOP	AL404447 T7 end of
c 42	44.4	2.7	110000	84	PFMAL13P2_1	Continuation (2 of
c 43	44.4	2.7	110000	84	PFMAL13P2_2	Continuation (3 of
c 44	44.2	2.7	82360	13	AP000606	AP000606 Arabidops
c 45	44.2	2.7	341064	75	AC074226	AC074226 Arabidops

ALIGNMENTS

RESULT	1
E12754	E12754
LOCUS	Gentianatriflora mRNA for acyltransferase,complete cds.
DEFINITION	24-JUN-1998
ACCESSION	E12754
VERSION	E12754.1 GI:3251586
KEYWORDS	JP 1997070290-A/2.
SOURCE	unidentified.
ORGANISM	unidentified.
REFERENCE	1 (bases 1 to 1622)
AUTHORS	Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y., Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL	Patent: JP 1997070290-A 2 18-MAR-1997;
COMMENT	SUNTORY LTD
OS	Gentianatriflora
PN	JP 1997070290-A/2
PD	18-MAR-1997
PF	30-JAN-1996 JP 1996046534
PI	17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
PC	ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO.
PI	FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC
PC	C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
CC	(C12N15/09,C12N15/19);
CC	strandedness: Double;
CC	topology: Linear;
CC	Key
CC	Location/Qualifiers
CC	Key

Db	781	AGCCACCTTCATCTACACCCCTATTGATATATCATGAAGCTCAAGAAATTCATTTCGTCAA	840	AUTHORS	Yonekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M.,
Qy	841	AAATCGCAACTTAAACGGGTAGTAGTAATATTAATCTGTCAACTTTTCACGGTGACATCTGC	900	TITLE	Fujiwara,H., Fukui,Y., Toshihiko,A., Yamaguchi,M. and Kusumi,T.
Db	841	AAATCGCAACTTAAACGGGTAGTAGTAATATTAATCTGTCAACTTTTCACGGTGACATCTGC	900		Molecular cloning and biochemical characterization of
Qy	901	ACTGATCTGGACATGCTTGTTCGAAATCATATTAGACACCGCTGTAAGAGAGAGGTTGGAAGA	960	JOURNAL	hydroxycinnamoyl-CoA:anthocyanin
Db	901	ACTGATCTGGACATGCTTGTTCGAAATCATATTAGACACCGCTGTAAGAGAGAGGTTGGAAGA	960	REFERENCE	3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla
Qy	961	GCATAAATCATGAGCAAACTTATGTGCTTTCATCACTGCGGCAACGCTTTTGCCTCCGCC	1020	AUTHORS	frutescens and diverse plant acyltransferase homologs
Db	961	GCATAAATCATGAGCAAACTTATGTGCTTTCATCACTGCGGCAACGCTTTTGCCTCCGCC	1020	TITLE	Unpublished (1999)
Qy	1021	GATACCTCAAAATTAATTTGGAAATTTGATAGTGCCTTGTATGTTGGTGGATCGACATCATGA	1080	JOURNAL	2 (bases 1 to 1622)
Db	1021	GATACCTCAAAATTAATTTGGAAATTTGATAGTGCCTTGTATGTTGGTGGATCGACATCATGA	1080	AUTHORS	Tanaka,Y. and Yonekura-Sakakibara,K.
Qy	1081	GCAACTTGTAGGAATGAAGGGTTCTCGGTAGCTGCAACCGCCATCGGAGATGCTATCCA	1140	TITLE	Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.
Db	1081	GCAACTTGTAGGAATGAAGGGTTCTCGGTAGCTGCAACCGCCATCGGAGATGCTATCCA	1140	JOURNAL	Yoshihiko Tanaka, Suntory Ltd., Institute for Fundamental Research;
Qy	1141	TAAGAGGTTACATGACTACGAAGGAATCTGAGAGAGATGCGATATCGCCGCCCGATC	1200		wakayazu dai 1-1-1, Shimamoto, Osaka 618-8503, Japan
Db	1141	TAAGAGGTTACATGACTACGAAGGAATCTGAGAGAGATGCGATATCGCCGCCCGATC	1200		(E-mail:Yoshihiko.Tanaka@suntory.co.jp, Tel:81-75-962-8807,
Qy	1201	AACATCTCGGCAACAGTCCGACGCTCATTTATCTGCTGGATCGGCACACGCAATGT	1260	FEATURES	Fax:81-75-962-8262)
Db	1201	AACATCTCGGCAACAGTCCGACGCTCATTTATCTGCTGGATCGGCACACGCAATGT	1260	source	1..1622
Qy	1261	GCATGATTTTGATGCAGATTTTGGTTGGGAAAGCTTGAAAAGCATGATCTGTTTCAAC	1320	gene	/organism="Gentiana triflora"
Db	1261	GCATGATTTTGATGCAGATTTTGGTTGGGAAAGCTTGAAAAGCATGATCTGTTTCAAC	1320	CDS	/db_xref="taxon:55190"
Qy	1321	TAATCCTCGGCAACACTAATTTTGCCTGCTCGGTCGAGAGATTTAAAGGAGCACTTGA	1380		35..1474
Db	1321	TAATCCTCGGCAACACTAATTTTGCCTGCTCGGTCGAGAGATTTAAAGGAGCACTTGA	1380		/gene="GAT106"
Qy	1381	GCTTGGCATTTCTTTCCTTGAAGATAGGATGACGACATTTGCGACCATTTTACGAATTT	1440		35..1474
Db	1381	GCTTGGCATTTCTTTCCTTGAAGATAGGATGACGACATTTGCGACCATTTTACGAATTT	1440		/gene="GAT106"
Qy	1441	CATCAATAGTCTCCATGTGAGGAGCCCTTGTGAAGAAAAAGTGTATCAATGTATAAAA	1500		/codon_start=1
Db	1441	CATCAATAGTCTCCATGTGAGGAGCCCTTGTGAAGAAAAAGTGTATCAATGTATAAAA	1500		/product="acyltransferase homolog"
Qy	1501	AGACAGACAAGTTATGATGCAACAATGTTTGTAGAGATTACAATCCATCGGGAAGATG	1560		/protein_id="BAA93452.1"
Db	1501	AGACAGACAAGTTATGATGCAACAATGTTTGTAGAGATTACAATCCATCGGGAAGATG	1560		/db_xref="GI:7415597"
Qy	1561	TATCAACTCATCTCTCTATATATATATTCAAATGTTTAAAAAATAAAAAA	1620		/translation="MAGSEDIKVKCRVAPPDVAEFTVPLSFFDMRWLISDAEH
Db	1561	TATCAACTCATCTCTCTATATATATATTCAAATGTTTAAAAAATAAAAAA	1620		HLHFRFRIPCNPKFISSKSLVLKHLPLAGNLIMPVDSDDRMPELVKKG
Qy	1621	AA 1622			SVSTIAEISSMDFYLAGDHQDSYKFNLDLPOLPEIVTSGDEVLPLAQVTVFSN
Db	1621	AA 1622			TGICIGRNLHVLKSGSFLHFNKLVLDKSNGLKFLPLSLSPMYDSSVVODPFH
RESULT	2				IRRTYNERKLLKQSGTPTVLNPKATSKDEVATFLLHPIDIMKLKKEFTSSKRNLTGS
AB026494	1622 bp	mRNA	PLN		SNYLSFTPTVSALITWCLSKSLDTVVREKVEEDKHAANLCATINCRORAPPYONY
LOCUS	Gentiana triflora GAT106	mRNA for acyltransferase homolog, complete cds	04-APR-2000		FGNCIVPCNVGSTEQLVGNBGLSVAATAIGDAIHKRHLDYEGILGRDWISPPRSTSA
DEFINITION	Gentiana triflora				APRSTLIYVGSQRNVHDFADFGKLEHESVSTNPSATLILISRRFRPKGALEL
ACCESSION	AB026494				GISLPKNRNDAPATFIETNFINSLHVRSPL"
VERSION	AB026494.1	GI:7415596			BASE COUNT 478 a 345 c 322 g 477 t
KEYWORDS	acyltransferase homolog.				Query Match 100.0%; Score 1622; DB 12; Length 1622;
SOURCE	Gentiana triflora				Best Local Similarity 100.0%; Pred. No. 0;
ORGANISM	Gentiana triflora				Matches 1622; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
REFERENCE	1 (sites)				Qy 1 GAACCATTCATCAATTAATCTGATTTAATTAAGATGGCAGGAAATTCGAGGATATCAA 60
					Db 1 GAACCATTCATCAATTAATCTGATTTAATTAAGATGGCAGGAAATTCGAGGATATCAA 60
					Qy 61 AGTTCTTGAGAAATGCCGTGTTGCGCCACACCGGACCGCGTCCGCGAGTTTACAGTCCC 120
					Db 61 AGTTCTTGAGAAATGCCGTGTTGCGCCACACCGGACCGCGTCCGCGAGTTTACAGTCCC 120
					Qy 121 ACTGTGTTTTTCGACATGCGATGGTTGATCTCTGATGAGAACACCATCTCGATTTCTTA 180
					Db 121 ACTGTGTTTTTCGACATGCGATGGTTGATCTCTGATGAGAACACCATCTCGATTTCTTA 180
					Qy 181 CAGATTCGCCCATCTTGTGCTCAACTCTAAATTTATCATTTTCATTAATAATCGTCCCT 240
					Db 181 CAGATTCGCCCATCTTGTGCTCAACTCTAAATTTATCATTTTCATTAATAATCGTCCCT 240
					Qy 241 TTCCCTTGTCTCAACACATTTCTTCCTAGCCGGAAATTTGATTTGCCCGGTAGATTTC 300
					Db 241 TTCCCTTGTCTCAACACATTTCTTCCTAGCCGGAAATTTGATTTGCCCGGTAGATTTC 300
					Qy 301 CTCCGATAGAAATGCCGGAGTTGCGTTTACAAGAAAGGGGACTCCGTTCTTTTACAAATTCG 360
					Db 301 CTCCGATAGAAATGCCGGAGTTGCGTTTACAAGAAAGGGGACTCCGTTCTTTTACAAATTCG 360
					Qy 361 AGAATCGAGCATGGATTTTGAATATCTCCCGGAGATCATCAGAGGATCTTATAAATTT 420
					Db 361 AGAATCGAGCATGGATTTTGAATATCTCCCGGAGATCATCAGAGGATCTTATAAATTT 420
					Qy 421 CAACGATTTGATTCGCGAGCTGCCAGAACCGATTGTGAACCTCCGCGGAGGAAGTATTACC 480

|||||
Db 421 CAACGATTTGATTCGCGAGTGCAGAACCGATTGTAACTCCGCGACGAAGTATTACC 480
Qy 481 ACTTTTGTCTTACAGGTGAGGGTGTCTCCAACACCGGTATATGCAATGGACCAATCT 540
Db 481 ACTTTTGTCTTACAGGTGAGGGTGTCTCCAACACCGGTATATGCAATGGACCAATCT 540
Qy 541 TCATCAAGTCTTGGTGATGCGACCTCTTTTCTGCAATTTTAAATAATATGGGTTTGGT 600
Db 541 TCATCAAGTCTTGGTGATGCGACCTCTTTTCTGCAATTTTAAATAATATGGGTTTGGT 600
Qy 601 TGCAAAATCCAAATCGAGATTCATTAAGTTCCTTCCACTTTCTTCTTACCTATGACGA 660
Db 601 TGCAAAATCCAAATCGAGATTCATTAAGTTCCTTCCACTTTCTTCTTACCTATGACGA 660
Qy 661 CAGATCTGTGTGCAAGATCCATTTCTATTCGTCGAAATCTACAAATGAAGAAACT 720
Db 661 CAGATCTGTGTGCAAGATCCATTTCTATTCGTCGAAATCTACAAATGAAGAAACT 720
Qy 721 GCTCAAAATCTCAGGCGACACCTACTGTCTTAAATCCAGCAATTTCTAAAGATGAAGTTCG 780
Db 721 GCTCAAAATCTCAGGCGACACCTACTGTCTTAAATCCAGCAATTTCTAAAGATGAAGTTCG 780
Qy 781 AGCCACCTTCACTTACACCTATGATGATCATGAAGCTCAAGAAATTCATTTGCTCAAA 840
Db 781 AGCCACCTTCACTTACACCTATGATGATCATGAAGCTCAAGAAATTCATTTGCTCAAA 840
Qy 841 AAATCGGAACCTTAACCGGTAGTACTTAATTAATCTGCACTTTCAGGTGACATCTGC 900
Db 841 AAATCGGAACCTTAACCGGTAGTACTTAATTAATCTGCACTTTCAGGTGACATCTGC 900
Qy 901 ACTGATCTGGACATGCTTGTGAAATCATTTAGACACCGTGTGAAGAGAGAGGTGGAAGA 960
Db 901 ACTGATCTGGACATGCTTGTGAAATCATTTAGACACCGTGTGAAGAGAGAGGTGGAAGA 960
Qy 961 GGATAACATCGACAACTTATGTGCTTTCATCACTGCGGACCAACGTTTGTCTCGCC 1020
Db 961 GGATAACATCGACAACTTATGTGCTTTCATCACTGCGGACCAACGTTTGTCTCGCC 1020
Qy 1021 GATACCTCAAAATTTACTTTGAAATTCGATGTCCTTGTATGCTGGATCGACTCATGA 1080
Db 1021 GATACCTCAAAATTTACTTTGAAATTCGATGTCCTTGTATGCTGGATCGACTCATGA 1080
Qy 1081 GCAACTTTGTAGAAATGAAGGTTTGTGCGTAGCTGCAACCGCCATCGAGATGCTATCCA 1140
Db 1081 GCAACTTTGTAGAAATGAAGGTTTGTGCGTAGCTGCAACCGCCATCGAGATGCTATCCA 1140
Qy 1141 TAAGAGTTTACATGACTACGAAGAAATCTGAGAGGAGATGGATATCGCGCCCGCATC 1200
Db 1141 TAAGAGTTTACATGACTACGAAGAAATCTGAGAGGAGATGGATATCGCGCCCGCATC 1200
Qy 1201 AACATCTGCGGCACCAAGGTGCGAGCTCATTTATGCTGGATCCGCACACCAATGT 1260
Db 1201 AACATCTGCGGCACCAAGGTGCGAGCTCATTTATGCTGGATCCGCACCAACCAATGT 1260
Qy 1261 GCATGATTTTGCATGACATTTGTTGGTGGGAAAGCTTGAAGAGCATGATCTGTTCAAC 1320
Db 1261 GCATGATTTTGCATGACATTTGTTGGTGGGAAAGCTTGAAGAGCATGATCTGTTCAAC 1320
Qy 1321 TAATCTTCCGCAACACTAAATTTTGTATCTCTCGGTCCAGAGATTTAAAGGAGCACTTGA 1380
Db 1321 TAATCTTCCGCAACACTAAATTTTGTATCTCTCGGTCCAGAGATTTAAAGGAGCACTTGA 1380
Qy 1381 GCTTGGCATTTCTTGGCTTAAGATAGATGGAGCGAATTTGCCACCAATTTTACGAATTT 1440
Db 1381 GCTTGGCATTTCTTGGCTTAAGATAGATGGAGCGAATTTGCCACCAATTTTACGAATTT 1440
Qy 1441 CATCAATAGTCTCCATGTGAGGAGCCCTTTGTAAGAAAAAGTGGTATCAATGTATATAA 1500
Db 1441 CATCAATAGTCTCCATGTGAGGAGCCCTTTGTAAGAAAAAGTGGTATCAATGTATATAA 1500
Qy 1501 AAGACAGACAAGTTATGATGCAACAAATGTTTATAGGAGATTACAAATCCATGCGAAGATG 1560
|||||

Db 1501 AAGACAGACAAGTTATGATGCAACAAATGTTTATAGAGATTACAAATCCATGCGAAGATG 1560
Qy 1561 TATCAAACTCATCTCTCTATATATATATATTTCAATGCTTTTAAAAAATAAAAAA 1620
Db 1561 TATCAAACTCATCTCTCTATATATATATATTTCAATGCTTTTAAAAAATAAAAAA 1620
Qy 1621 AA 1622
Db 1621 AA 1622
RESULT 3
E12756
LOCUS
DEFINITION
Perilla ocimoides mRNA for acyltransferase, partial cds.
ACCESSION
E12756
VERSION
E12756.1 GI:3251588
KEYWORDS
JP 1997070290-A/4
SOURCE
unidentified.
ORGANISM
unclassified.
REFERENCE
1 (bases 1 to 1479)
AUTHORS
Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y.,
Yonekura, K., Mizutani, M. and Kusumi, T.
TITLE
GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL
Patent: JP 1997070290-A 4 18-MAR-1997;
SUNTORY LTD
COMMENT
OS Perilla ocimoides
PN JP 1997070290-A/4
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95p 67159, 29-JUN-1995 JP 95p 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
FUJIKU YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAOKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FH source 1..1479
FT /organism='Perilla ocimoides' FT
FT /tissue_type='leaves',
FT /clone='pSAT208'
FT CDS 1..1343
FT /product='acyltransferase'.
FEATURES
source Location/Qualifiers
1..1479 /organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 420 a 316 c 331 g 411 t 1 others
ORIGIN
Query Match 13.4%; Score 218; DB 10; Length 1479;
Best Local Similarity 50.8%; Pred. No. 4.6e-44;
Matches 741; Conservative 1; Mismatches 661; Indels 57; Gaps 7;
Qy 62 GTTCTTGGAGAAATGCCGTGTTCGCCACACCGCGCGCTGCGGAGTTTACAGTCCCA 121
Db 3 GTGATCGAAACGTGTAGAGTTGGCCGCCGCCGACTCGGTGGCGGAGCAATCGGTCCG 62
Qy 122 CTGTCGTTTTTGCATGCGGATGGTGTGATCTCTGATGCGAAGACCACTATGCAATTTCTAC 181
Db 63 CTCACATTTCTCGACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCCTCTTCTAC 122
Qy 182 AGATTCGCCCATCTTCTGCCAAC--TCTAAATTTATCATTTTCATCATTAATTCGTC 238
Db 123 GAATTCCTTGTTCGCAAGCAACATTTTTCAGATTCATTCGTTCCAAACATCAACAACT 182
Qy 239 CTTTCCCTTGTCTCAACACACTTTTCTTCCGTTAGCCGGGAATTTGATTGGCCGCTAGAT 298

Db 63 CTCACATCTTCGACATGACGTGCTGCAATTTTCATCCCATGCTTCAGCTCCTCTCTAC 122
Qy 182 AGATTCCGCCATCTCTGTCGCCAAC---TCTAAATTTATCATTTTCATCCATTAATAATCGTCC 238
Db 123 GAATTCCTCTGTCGAAGCAACATTTTTCAGATCCATCTGTTCCAAACTCAAAATCT 182
Qy 239 CTTTCCCTCTGTCGAACACTCTTCTCCGTTAGCCGGAATTTGATTTGGCCGCTAGAT 298
Db 183 CTCCTAAACCTCATACACTCTCTCCCTCTCTCATGCAATTTAACTACCCCTCA--- 239
Qy 299 TCTCCGATAGATCCCGAGTTCGGTTCAAGAAAGGGGACTCCGTTTCTTTTAAACAT 358
Db 240 TCCCGGAGAAATGCCGAGTTTCGGTATCTATCCGCGGACTCGGTTTCTTTACCATC 299
Qy 359 GCAGATCGAGCATGGATTTGATATCTCCGCGGAGATCATCAGAGGATTTCTTATAA 418
Db 300 CGAATCTACGAGGACTTCGATGATCTGTCGGAATCTGTCAGAACTCCTCCGTTAGG 359
Qy 419 TTCAACGATTTGATTCGCGAGTCCGAGAACCGATTTGTAACCTCCGCGGAGCAAGTATTA 478
Db 360 CTCACAACTTTGTCCTTAATTTGCCGCCATTTGTCGAAGAATCCGATAGAAACTCTTC 419
Qy 479 CCACTTTTGTCTTACAGTCAAGTGTCTTCCACACACCGGTATATGATTTGGAGCGAAT 538
Db 420 CAAGTTTTCGCGGTGCGAGTGACTCTTTTCCAGCGCGAGCGTGGTATTTGGAATAGCA 479
Qy 539 CTTTCATCAAGTCTTGTGTGATGTCAGTCTTCTTCTGCAATTTAATAAATATGAGTTTG 598
Db 480 ACGCATCACAGTTGAGGACCGCGCTGTTCTTCGGGTTTATAACGGTTCGTCTTCA 539
Qy 599 GTTGACAAATCCAATGGAGATTTCAATAAAGTTCTTCCACATTTCTTCTACCTATGTAC 658
Db 540 ATGACCAACACATTTGAATGAAGATCAAGATCAAGAAATTTAATCTTCCAGATTTC 599
Qy 659 GACAGATCTGTTGTCGAAGATTCATTTCAATTCGTCGAAGAAATTTCAAAATGAAGAAA 718
Db 600 GATAGATCCGTGATATAATATCCGACGAAATTT-----GACTCCATTTATTTGGAGAAC 653
Qy 719 CTGCTCAATCTCAGGCGACACCTACTCTTCTTAATCCAGCAATTTCTAAGATGAAGTT 778
Db 654 CGCTAAATTTCTTTGCAATCTCGT-----CATCCTCATACCGACGCGCGGAT 707
Qy 779 CGAGCCACTTTCATCCATACACCTTATGATATCATGAGCTCAAGAAATTCATTTCCGTCA 838
Db 708 CGAACCAGTTTCGTTTACCCCAATCCAAATTAAGAAATTAAGGTTTGGATTCAGTCC 767
Qy 839 AAAATCCCAACTTAACCGGTAGTAGTAATTAATCTGTCAACTTTTCCAGCGTCACTCT 898
Db 768 AGAGTTCCAAAGTTTAGTC-----CATCTCTCATCTTTTGTAGCGATTGCA 812
Qy 899 GCACTGTCTCGACATGCTTCTGCAATCATTTAGACACCGCTGTAAGAGAGAGAGTGAA 958
Db 813 CTTATATGTCGGTGGCATTAACGAATCATC-----ACAGCAGATGAAGAC 860
Qy 959 GAGGATAAATGACGACAACTTATGTCTTTCATCACTCCGACACAGTTTGTCTCCG 1018
Db 861 CAAGACACGAGGATGCAATTTTCTTGATTCGGTCTGATCTAAGCCACGATTAGATCCG 920
Qy 1019 CCGATACCTCAAAATTCATTTGGAAATTCATAGTGCCTGTATGTTGGGATTCAGCTCAT 1078
Db 921 CCGTTCTCTGAAATTAATCTTCGGAACTGCTTATCTGACGCGCTGCCGAGAAATCGCGG 980
Qy 1079 GAGCAACTTGTAGGAATGAAGGTTGTGCTGCTAGCTCAACCGCATCGGAGATGCTATC 1138
Db 981 CGAGAGCTGGGAGAGAAAGGGTGTCTTCGACCTGAGGTAATTCGCGCGCGAGATA 1040
Qy 1139 CATAAGAGTTTACATGACTACGAGGAATTTCTGAGGAGATTTGGATATCGCCGCCCGCA 1198
Db 1041 AAAAAAGGATCAACGACAGAGAAATTTAGAAACGGTGGAGAAATGTCGCCCGGAGATT 1100
Qy 1199 TCAACATCTCGCGGACCAAGTCCACGCTCATTTATGCTGTTGATCGGCACACGCAAT 1258

Db 1101 CGTAAAGC-----GTTGCAGAAATCATATTTTTCGGTGGCAGGATCGAGCAAG 1148
Qy 1259 GTGCATGATTTTGATGCAGATTTTGGTGGGAAAGCTTGAAGACGATGATCTGTTTCA 1318
Db 1149 CTAGATCTTTACGTTGCAGATTTTGGATGGGAGGCGAGCAAGCAAGAAATATTGTCG 1208
Qy 1319 ACTAATCCTTCGCGAACACTAAATTTTGATCTCTCGTCCGTCACAGAAGATTTAAAGGACACTT 1378
Db 1209 ATTGATGGGAGAAATATGCAATGACACTTTTGAAGCCAGGATTTTGAAGAGGAGATTG 1268
Qy 1379 GAGCTTGGCATTTCTTTCCTTAAAGATAGGATGACGACATTTGCCACCATTTTACGAA 1438
Db 1269 GAGTTTGTCTTCTTTCCTTAAAGCAAAATGATGCTTTTGTCTTATTTTTCAGTG 1328
Qy 1439 TTCATCAATAGT 1450
Db 1329 GGAATTAATGTT 1340

RESULT 5
AB010708 LOCUS 1679 bp mRNA PLN 20-FEB-1999
DEFINITION Gentiana triflora mRNA for Anthocyanin 5-aromatic acyltransferase,
complete cds.
ACCESSION AB010708
VERSION AB010708.1 GI:4185598
KEYWORDS Anthocyanin 5-aromatic acyltransferase.
SOURCE Gentiana triflora petal cDNA to mRNA, clone: pCAT4.
ORGANISM Gentiana triflora
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
1 (bases 1 to 1679)
Fujiwara,H., Tanaka,Y. and Kusumi,T.
Direct Submission
Submitted (22-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Hiroaki Fujiwara, Sutoryi LTD., Inst. of Fundamental Research;
1-1-1, Wakayamadai, Shamoto-cho, Osaka 618-0024, Japan
(E-mail: Hiroaki.Fujiwara@sutoryi.co.jp, Tel: +81-75-962-8807,
Fax: +81-75-962-8262)
2 (sites)
Fujiwara,H., Tanaka,Y., Yonekura-Sakakibara,K.,
Fukuchi-Mizutani,M., Nakao,M., Fukui,Y., Yamaguchi,M., Ashikari,T.
and Kusumi,T.
cDNA cloning, gene expression and subcellular localization of
anthocyanin 5-aromatic acyltransferase from Gentiana triflora
Plant J. 16 (4), 421-431 (1998)
99097837 Location/Qualifiers
1. 1679
/organism="Gentiana triflora"
/db_xref="taxon:55190"
/clone="pCAT4"
/tissue_type="petal"
6. 1415
/codon_start=1
/product="Anthocyanin 5-aromatic acyltransferase"
/protein_id="BAA74428.1"
/db_xref="GI:4185598"
/translation="MEIQMKVLEKCVTPPSDPTDVELSLPVTTFDIPWLHLNKMQ
SLFLYDFPYPRTHFLDTPVLPNKLASLTLKHYPLVLSGNLIMPDKSGEMPKFQYSRDE
GDSITITVAESDQDFYLGKQLVDSNDLHGLFVMPRVITMDQYKVIPLVAQVTV
FPMRGIAVALTAHSDAKSFVMFNANAY INKFRDADLLSANLPSFDRSTIKDL
YGLEETFWNEMODVLEMFSGSPKPPKVRATYVLSLAETOKLNKVLNLRGSEPT
IRVTTFTMTCGYVWTCMVSKDDVSESSNDENELEYFSFTACRGLLTTPPCPNPF
GACLSCVAKATKELVGDLLVAVAAGEATEKRLHNEKFLADAKTWLSESNGLP
SKFLGITGSPKEDSGVDFGKPKAFDITSDYAEILYVIOSRDFKGVETGVSILP
1679
/notes="24 a nucleotides"
polyA_site 488 a 350 c 359 g 482 t
BASE COUNT
ORIGIN

Query Match		11.7%;	Score 189.2;	DB 12;	Length 1679;
Best Local Similarity		52.2%;	Pred. No. 7.8e-37;		
Matches 728;		Conservative 0;	Mismatches 618;	Indels 48;	Gaps 12;
QY	57	TCAAGTCTTCTGAGAAATCCGTTGTCGCCACACCGGAC---	CGCGTCGCCGAGTTTA	113	
DB	25	TGAAGTCTTCTGAAATGCCAAGTTACACCACTCTGACACACAGATGTCGAGTTAT	84		
QY	114	CAGTCCCACTGCTGTTTTCGACATCGGATGGTTGATCTCTGATCGAGAACACCATCTGC	173		
DB	85	CGCTACCGGTAACATCTTTCGATATCCCTGGTTGCACCTTGAATAAGATGCGAGTCCCTTC	144		
QY	174	ATTCTACAGATTCCGCCATCTCTGTGCCAA---CFTAAATTTATCATTTTCATCCATTA	230		
DB	145	TGTTTTACGACTTTCGTAACCAAGAACACATTTCTTGGACACTGTATCCCTTAATCTTA	204		
QY	231	AATCGTCCCTTTCCCTTGTCTCAACACTTCTTCCGTTAGCCGGAAATTTGATTTGCG	290		
DB	205	AGGCCTCTTTGCTCTCACTCTAAACACTAGCTTCCGCTTAGCGGAAATTTGTTGATGC	264		
QY	291	CGGTAGATTCTCCCGATAGAAATGCCGAGCTTGGTTAC-----AGAAGGGGACTCG	344		
DB	265	CGATCAAAATCGGGCGA---AATGCCGAAGTTTCACTACTCCCGTATGAGGGGACTCGA	321		
QY	345	TTCTTTTAAACAATTGCGAATCGAGCATGGATTTGATTTATCTCTCGCCGGAGATCATCAGA	404		
DB	322	TAACTTTGATCGTTGGGAGCTGACACGAGATTTTGACACTACCTTAAAGGTCACTCAACTGG	381		
QY	405	GGGATTTCTTATAAATTAACAGATTTTGATTTCCGAGCTGCGCAGAACCGATTTGTAACC---	461		
DB	382	TAGATTCCCAATGATTTGCGCTCTTTTATGTTATGCCACGGGTTATAAGGACCATGC	441		
QY	462	CGGGCACCAGATTAATACCACTTTTTCCTTTTACAGTGGAGGTGTTCTCCACACCGGTA	521		
DB	442	AAGACTATAAAGTGATCCCGCTCGTAGCCGTGCAAGTAACCGTTTTCCTAACCCGTGGCA	501		
QY	522	TATGCAATTTGGACGCAATCTTTCATCAAGTTCTTGGTGCATGCCAGTCTTTTCTGCAATTTTA	581		
DB	502	TAGCCGTGCTCTGACGGCACATCATTTCAATTCGACATGCTAAAGATTTTGTAAATGTTCA	561		
QY	582	ATAAATATGGGTTTGGTTGACAATCCAATGAGGATTCATTAAGTTTCTTCCACTTT	641		
DB	562	TCAATGCTTGGGCTATATAACAAT---TTGGGAAGACGGGACTTGTGTCGCGCA	618		
QY	642	CTTCTCTACTATGATACGACAGATCTGGTGCAAGATCCATTTTCATATCTCGTGAATAA	701		
DB	619	ATCTTCTTCCATCTTTCGATAGATGATAATCAAGATCTGTATGCCCTAGAGGAAACAT	678		
QY	702	TCTACAATGAAGAAACTGCTCAAAATCTCAGGGCACACTACTGTCTTAAATCCAGCAA	761		
DB	679	TTTGGAACCAATGCAAGATGT---TCTTGAATGTCTCTAGATTGGAAGCAACCCC	735		
QY	762	TTTCTAAAGATGAAGTTCGAGCACCTTCATCTACACCCCTATTTGATATCATGAAGCTCA	821		
DB	736	CTCGATTCAACAAGGTACGAGCTACATATGTCTCTCCCTTGTGTAATCCAGAAGCTAA	795		
QY	822	AGAAATTCATTTTCGTCAAAAATCGCAACTTAACCGGTAGTAGTAATTTATAATCTGCAA	881		
DB	796	AG-AACAAAGTACTGAATCTCAGAGGATCCGAACCCAGCAATA-----CGGTAAACGA	846		
QY	882	CTTTCACGGTGACATCTGCACTGATCTGACATGCTTGTGCAANTCATTTAGAACCCGTCG	941		
DB	847	CGTTCAANTGACGCTGGATAGCTATGACATGCTGATGTCATGTCATTAAGATGACGCTCG	906		
QY	942	TAAAGAGAA---GGTGAAGAGGATAAATATGACAGCAAACTTATGTCTTTTCATCAACT	998		
DB	907	TATCAGAGGAATCATCAACGAGCAAAATGAGCTCGAGTACTTCACTTTTACAGCGGAT	966		
QY	999	GCGGACAACTTTTGTCTCGCCGATACCTCAAAATTAATCTTGTGAATTTGATAGTGCCTT	1058		
DB	967	GCGGAGGACTTCTGAGCGCCCGCTGTCCGCTAACTACTTTTGGCAACTGTCTTGGCTCAT	1026		

QY	1059	GTATGCTGGGATCGACTCATGAGCAACTTGTAGGAATGAAGGTTTTCGGTAGCTGCAA	1118		
DB	1027	CGCTTCCAAAACCAACACATAAAGAGTTAGTTGGGATFAAAGGCTTCTTGTTCAGTTG	1086		
QY	1119	CGCCCATCGGAGATGCTATCCATTAAGAGGTTTACATCACTACGAAAGAAATTTCTGAGAGG	1178		
DB	1087	CAGCTATTGGAGAAGCATTTGAAAAGAGGTTGCACACGAAAAAGGGTCTTCTTCAGATG	1146		
QY	1179	ATTGGATATCGCGCCCGCCGNTCAACATCTCGCGCACCAAGGTCGACGCTCATTTATGTCG	1238		
DB	1147	CAAAAACCTTGGTTATCGGAATCTAAT-----GGAATCCCTTCAAAAAGATTTCTCG	1197		
QY	1239	TTGGATCCCAACCAACGCAATGTGCATGATTTTGTGATCGATTTTGGTGGGAAAGCTTG	1298		
DB	1198	GGATTACCGGATCGCCTAAGTTCGATTCGTATGTTGATTTTGGATGGGAAAGCCTG	1257		
QY	1299	AAAGCATGAATCTGTTTCACTAATCCTTTCGGCAACACTAATTTTGATCTCTCGGTCCA	1358		
DB	1258	CAAAATTTGACA---TTAGCTCTGTTGATTATGCAGAAATTTGATTATGTTGATTCAGTCCA	1314		
QY	1359	GAAGATTTAAAGGAGCACTTGAGCTTGGCATTTTTCCTTGCCTAAGAAATAGGATGACGCAT	1418		
DB	1315	GGATTTTGAAGAAGGTGCGAGATTGGAGTATCATTTGCCCTAAGATTCATATGGATGCAT	1374		
QY	1419	TTGCCACCACTTTT 1432			
DB	1375	TTGCAAAATCTTT 1388			
RESULT	6				
LOCUS	E12753	1703 bp	DNA	PAT	24-JUN-1998
DEFINITION	Gentianatriflora mRNA acyltransferase, complete cds.				
ACCESSION	E12753				
VERSION	E12753.1	GI:3251585			
KEYWORDS	JP 1997070290-A/1.				
SOURCE	unidentified.				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 1703)				
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.				
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY				
JOURNAL	Patent: JP 1997070290-A 1 18-MAR-1997;				
COMMENT	SUNTORY LTD				
	OS	Gentianatriflora			
	PN	JP 1997070290-A/1			
	PD	18-MAR-1997			
	PF	30-JAN-1996	JP 1996046534		
	PR	17-FEB-1995	JP 95P 67159, 29-JUN-1995	JP 95P 196915	PI
	ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO				
	MASHIRO, FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAHIKI				
	PC C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC				
	C12R1:865),				
	PC	(C12N9/10, C12R1:19);			
	CC	Strandedness: Double;			
	CC	Topology: Linear;			
	PH	Key	Location/Qualifiers		
	FT	source	1. .1703		
	FT		/organism='Gentianatriflora'		
	FT		/variety='japonica'		
	FT		/tissue_type='petal'		
	FT		/clone='pgAT4'		
	FT	CDS	6. .1415		
	FT		/product='acyltransferase precursor'	FT	
	mat_peptide	9. .1412	/product='acyltransferase'		
FEATURES	Location/Qualifiers				
	1. .1703				
	/organism='unidentified'				
	/db_xref='taxon:32644'				

BASE COUNT 512 a 350 c 359 g 482 t

Query Match 11.7%; Score 189.2; DB 10; Length 1703;
Best Local Similarity 52.2%; Pred. No. 7.9e-37;
Matches 728; Conservative 0; Mismatches 618; Indels 48; Gaps 12;

QY 57 TCAAGTCTTGGAGAAATGCGGTGTCGGCCACCACCGGAC---GCCGTGCGCCGAGTTTA 113
DB 25 TGAAGTCTTGAANAATGCCAAGTTACACCACCATCTGCACACACAGATGTCGAGTTAT 84
QY 114 CAGTCCACATGTCGCTTTTTCGACATGCGATGGTTGATCTCTGATGCGAGAACACCATCTGC 173
DB 85 CGTACCGGTAAACATCTTCGATATCCCTGGTTGCACTTGAATAAGATGCAAGTCCCTTC 144
QY 174 ATTTCTACAGATTCGGCCATCTTGTCCCAA---CTCTAAATTTATCATTTTCATCCATTA 230
DB 145 TGTTTTACGACTTTCCGTACCCAAAGAACACATTTCTTGGACACTGTTATCCCTAATCTTA 204
QY 231 AATCGTCCCTTTCCTTGTCTCAACACATTTCTTCCGTTAGCCGGGAATTTGATTTGGC 290
DB 205 AGGCTCTTGTCTCTCACTCTAAACACATAGTTCGCTTAGCGGAATTTGTTGATGC 264
QY 291 CGTATGATTCCTCCGATAGATGATGATGCGGAGTTGCGTTAC-----AAGAAAGGGGACTCCG 344
DB 265 CGATCAATCGGGCGA---AATCGGAAGTTTTCAGTACTCCCGTGTAGGGCGGACTCGA 321
QY 345 TTTCTTTAACAATTTGCAGATCGAGATGATTTTGTATTTATCTCCCGGAGATCATCAGA 404
DB 322 TAATCTTTGATGTTGCGGAGTCTGACGAGGATTTTGTACTTAAAGGTCATCAACTGG 381
QY 405 GGGATTTCTTATAAATTTCAACGATTTGATTCGCGAGTCCGACAGACCGATTTGTAACC---T 461
DB 382 TAGATTTCAATGATTTGCATGGCTTTTATGTTATGTCACCGGTTATTAAGGACCATGC 441
QY 462 CCGCGCAGAGATTAATACCACTTTTGTCTTACAGGTGACGGTGTCTTCCCAACACCGGTA 521
DB 442 AAGACTATAAGTATCCGCTCGTAGCGGTGCAAGTAACCGTTTTCCTTAACCGTGCGA 501
QY 522 TATGATTTGGAGCAATTTTCATCAAGTCTTGTGTGATGCGAGTCTTCTTCTGCAATTTA 581
DB 502 TAGCGGTGGTCTTACGGCATCATTTCAATTTGCAGATGCTTAAAGTTTGTATGTTCA 561
QY 582 ATAAATATGTTGTTGTTGACAAATCCAAATGAGATTTCAATTAAGTTCCTTCCACTTT 641
DB 562 TCAATGCTTGGCCCTATATTAACAAAT---TTGGGAAGACCGGACTTGTGTCCCGCA 618
QY 642 CTTCTACCTATGACACAGATCTGTGGTGCAAGATTCATTTATTCGTGCAAAAA 701
DB 619 ATCTTCTTCCATCTTTCGATAGATCGATAATCAAGATCTGTATGGCTAGAGGAACAT 678
QY 702 TCTACATGAAGAAACTGCTCAATCTCAGGCGACACCTACTGTTCTTAATCCAGCA 761
DB 679 TTTGGAACGAATGCAAGATGT---TCTTGAATTTGTTCTAGATTTTGGAGCAAAACCC 735
QY 762 TTTCTAAAGATGAAGTTTCGAGCCACCTTTCATCCCTACACCTTATGATATCATGAAGTCA 821
DB 736 CTCGATTTCAACAAGTAGGAGCTACATATGCTCTCTCCCTTGTGAATCCAGAGCTAA 795
QY 822 AGAATTTATTTCTGCAAAAAATCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA 881
DB 796 AG-AACAAAGTACTGTAATCTCAGAGGATCCGACCGACAATA-----CGTGTAAACA 846
QY 882 CTTTACGGGTGACATCTGCATGATCTGGACATGCTTGTGAAATCATTAGACACCGTCG 941
DB 847 CGTTCAAAATGACGTGGTATGATGATGATGATGATGATGATGATGATGATGATGATG 906
QY 942 TAAGAGAGAA---GGTGAAGAGGATTAACATGACGCAAACTTATGTGCTTTTCAACT 998
DB 907 TATCAGAGGAATCATCGAACGACGAAATAGAGCTCGAGTACTTTCAGTTTACACGGAT 966
QY 999 GCCGACAAAGTTTGTCTCCCGCGATACCTCAAAATTAATTTGGAATTTGCAATGATGCTT 1058

DB 967 GCCGAGGACTTCTGACGCCGCCGCTGTCGCTAACTACTTTGGCAACTGTCTTCGTCAT 1026
QY 1059 GTATGTTGGATCGACTCATGAGCAACTTGTAGCAAAATGAAGGTTGTCTCGTAGCTCAA 1118
DB 1027 GCGTTGCAAAAGCAACACATAAAGAGTTAGTTGGGATAAAAGGCTTCTTGTTCAGTTG 1086
QY 1119 CCGCATCGGAGATGCTATCATTAAGAGGTTACATGACTACGAAGAAATTTCTGAGAGAG 1178
DB 1087 CAGCTATTGGAGAGCCATTGAAAGAGTTGCAACGAAAGAGGCTTCTTTCAGATG 1146
QY 1179 ATTGGATATCGCCGCCCGATCAACATCTGCGGCAACCAAGTGCAGCGCTCATTTATGTCG 1238
DB 1147 CAAAAAATCTGTTTATCGGAATCTAAT-----GGAATCCCTTCAAAAAGATTTCTCG 1197
QY 1239 TTGATCGGCACACGCAATCTGCATGATTTTGATGACAGATTTTGGTGGGAAAGCTTG 1298
DB 1198 GGATTCGCGATCGCTTAAGTTCGATTCGATGTTGATGATTTTGGATGGGAAAGCGTG 1257
QY 1299 AAAAGCATGAATCTGTTTCAACTAATCTTCGCAACACTAATTTTGATCTCTCGGTCCA 1358
DB 1258 CAAATTTGACA---TTACCTCTGTTGATGATGAGATTTGATGATGATGATGATGATG 1314
QY 1359 GAAGATTTAAAGAGGACCTTGAAGTTCGCAATTTCTTTCCTTAAGATAGGATGAGCGCAT 1418
DB 1315 GGGATTTTGAAGAGGTTGGAGATTTGGAGTATGATGATGATGATGATGATGATGATGATG 1374
QY 1419 TTGCCACCAATTTT 1432
DB 1375 TTGCAAAATCTTT 1388

RESULT 7
E12757
LOCUS E12757 1508 bp DNA PAT 24-JUN-1998
DEFINITION Senecio cruentus mRNA for acyltransferase,,partial cds.
ACCESSION E12757
VERSION E12757.1 GI:3251589
KEYWORDS JP 1997070290-A/5.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 1508)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 5 18-MAR-1997;
SUNTORY LTD
COMMENT OS Senecio cruentus
PN JP 1997070290-A/5
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAARI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FT source 1..1508
FT /organism='Senecio cruentus'
FT /tissue_type='petal.'
FT /clone='pCAT48'
FT CDS 1..1367
FT /product='acyltransferase'.
FEATURES Location/Qualifiers
source 1..1508
/organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 442 a 293 c 296 g 477 t
ORIGIN

Query Match 8.6%; Score 139.8; DB 10; Length 1508;
Best Local Similarity 48.7%; Pred. No. 1.9e-24;
Matches 677; Conservative 0; Mismatches 642; Indels 72; Gaps 8;

QY 57 TCAAGCTTCTGAGAAATCCGCTGTTGCGCCACACCGGCGCGGAGTTTACAG 116
DB 1 TGAACATCTCGAACATGCGCCGAATATCGGCCCTCGGCGACCATCGGCCATCGCTCGT 60
QY 117 TCCCACTGCTGTTTTCGACATCGCATGTGTATCTCTGATGAGAACACCATCTGCATT 176
DB 61 TATCTCTTACTTCTTCGACATTAATCTGCTACTCTCTCCCTCCGGTCCACCATCTTTTCT 120
QY 177 TCTACAGATTCGCGCATCTCTGTCGCAACTCTAAATTTATCATTTTCATCCCA---TTAAAT 233
DB 121 TCTATGACTTTCACATCTCTAAATCCCATTTTCATGGACACTAATTTGTTCCGAGGCTAAAC 180
QY 234 CGTCCCTTTCCCTGTCTCAACACTTTCTTCCGTTAGCGGGAATTTGATTTGGCCGG 293
DB 181 AATCTTTATCGGTCACTCTCAACATTTTTCGGTTTGTAGTAATTTGATTTGATTTTC 240
QY 294 TAGATTCCTCG-----ATAGATGCCGGAGTTGCGTTTACAAGAAAGGGGACT 341
DB 241 CTAACACTGATGTTGCGGTTTAAATAAACCAGAAATAAACACGTTTGAAGGTGATT 300
QY 342 CGGTTTCTTAACAATTCGAGATCAGCATGGATTTTGTATATCTCGCGGAGATCATC 401
DB 301 CTGTTGTTGTTACTTTTGCAGAAATGTTGCTTGACTTTTAAATAATTTGACAGGAATCATC 360
QY 402 AGAGGATTTCTTAATAATCAACGATTTTGATTCGCGAGCTGCCAGAACGATTTGAACCT 461
DB 361 CTCGAAATGTGAAACTTTTATCCACTTGTACCTTCATTTGGGAATGCAATCAAAATTA 420
QY 462 CGCGGACGCAAGTATTTACCACCTTTTTCGTTTACAGGTGACGGTGTCTTCCAAACCGGTA 521
DB 421 GTGATTCGCGACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTTTCGGGCTCGGTA 480
QY 522 TATGCAATTCGACCAATCTTCATCAAGTCTCTGGTGATGCGACTTCTTTCTGCATTTTA 581
DB 481 TATCACTAGGAATGACGAATCATCATAGCTTGGTGACGCTAGCAGCGGTTCAACTTTT 540
QY 582 ATAAATATGCGGTTTGGTTGCAAAATCCAATCGACATTCATTAAGTTCCTTCCACTTT 641
DB 541 TGAAGGGTGGACTTCGATTTTCAATCTGGTGTAGT-----CGGT 582
QY 642 CTCTCTACTATGTACGACAGATCTGTGTGCAAGATCCATTTTCATATTCGTCGAAAA 701
DB 583 CTTTTTTAACGAAAGGATCTCCACCGTTTTCATAGATTTATACATCCACATTTAG 642
QY 702 TCTACAATGAAGAAATCTGCTCAATCTCAGGGCACACTACTGTCTAAATPCCAGCAA 761
DB 643 ATGAAATTAAGTTGAGACATACAAAGCTCGAAAGTTTATATAACCTTCGAGCCTGTG 702
QY 762 TTCTAAAGATGAAGTTCGAGCACCTTCATCTACACCTTATGATATCATCAAGCTCA 821
DB 703 GTCCCACTGATAAGTTGCGTCAACCTTTGTTGACCCGAACTAATATCAATCTACTAA 762
QY 822 AGAAATTCATTTGCGTCAAAAAATTCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA 881
DB 763 AGAAAGGTCTTAACCAAGTGCCAACTT-----GGAGTACATGTCAT 807
QY 882 CTTTCAGGTGACATCTGCACATGATCTGGACATCTGTGCGAAATCATTTAGACACCGTCG 941
DB 808 CTTTACGGTAACCTTGTGTTATATATATGAGTTGATAGCAAAATCACTCGTAAAAATAG 867
QY 942 TAGAGAGAGGTGGAAGAGGATAACATGTCAGCAAACTTATGTGCTTCATCACTGCC 1001
DB 868 GAGAAAGAAAGGGCGAAGACG-----AGTTAGAACAGTTTCAATCACTATTTGTTGTC 921
QY 1002 GACAAGGTTTGTCTCCGCGGATACCTCAAAATTAATTTGGAATTTGCATAGTCCCTTGTA 1061

DB 922 GATCTCGTCTTGATCCACCAATTCACCAGCCTACTTTGGTAACTGTGGTGCCACCATGTG 981
QY 1062 TGGTGGATCGCATCATGAGCAACTTTGTAGGAAATGAAGGTTTGTGGTAGCTGCACCG 1121
DB 982 TCCCGACCTTAAAAAATGTCGTTTGTGACTAGCGAAAATGGGTATGCACTTGGTGCTAAAG 1041
QY 1122 CCATCGGAGATGCTATCCATAAGAGGTTTACATGACTACGAGGAATTTCTCAGAGGAGATT 1181
DB 1042 TAATTCGAGAGTCTATATGCAAAATGATATATTAATAGGACGGAATCTTGAAGATGCCG 1101
QY 1182 GGATATCGCGCCCGCCGATCAACATCTCGGCGACCAAGGTGCGACGCTCATTTATGTGTTG 1241
DB 1102 CGAGATGGCATGAACCTTTCATGATCCCG-----CTAGGAAGATTGGTGTGCTG 1152
QY 1242 GATCCGACACACCAATGTCATGATTTTGTATGATGATTTTGGTGGGAAAGCTGAAA 1301
DB 1153 GTACACTAAGCTCAACTTGTACGACTTT-----GATTTTGGTGGGGAAGCCGATAA 1206
QY 1302 AGCATGAATCTGTTTCAACTTAATCTTCCGCAACACTAATTTTGTATCTCTCGGTCCAGAA 1361
DB 1207 AGTATGAGACTGTTTCAATAGA---CTATATAGCTCGATTTCTATATAATGCAAGCAAAA 1263
QY 1362 GATTTAAAGGAGCATTGAGCTTGGCATTTCTTTGCTAAGAAATAGGATGGACCATTTG 1421
DB 1264 CATCAGCACAAAGATCTTGAAATTTGATGCTACCGAGTATGCAAAATGGAGCGGTTTT 1323
QY 1422 CCACCATTTT 1432
DB 1324 CTAGCATCTTT 1334

RESULT 8
AC005308/c

LOCUS AC005308 178273 bp DNA HTG 12-AUG-2000
DEFINITION Plasmodium falciparum chromosome 12 clone 3D7, *** SEQUENCING IN
PROGRESS ***, 1 ordered pieces.

ACCESSION AC005308
VERSION AC005308.8 GI:9797719
KEYWORDS HTG; HTGS_PHASE2.
SOURCE malaria parasite P. falciparum.
ORGANISM Plasmodium falciparum
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.

REFERENCE 1 (bases 1 to 178273)
Hymn, R.W., Fung, E.L., Qin, F., Rowley, D., Mao, J., Tamaki, T.,
Kurd, O.B., Conway, A.B. and Davis, R.W.
Plasmodium falciparum 3D7 chromosome 12
Unpublished

TITLE DIRECT SUBMISSION
REFERENCE 2 (bases 1 to 173273)
Hymn, R.W., Qin, F., Fung, E.L., Conway, A.B. and Davis, R.W.
AUTHORS Direct Submission
JOURNAL Submitted (21-JUL-1998) Stanford DNA Sequencing and Technology
Center, Stanford University, 855 California Avenue, Palo Alto, CA
94304, USA

COMMENT On Aug 12, 2000 this sequence version replaced gi:8810450.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 1 178273: contig of 178273 bp in length.

FEATURES
source Location/Qualifiers
1. 178273
/organism="Plasmodium falciparum"
/db_xref="taxon:5833"
/chromosome="12"
/clone="RFAAC492"
/clone="3D7"
BASE COUNT 71271 a 17228 c 17167 g 73607 t

Qy	1476	AAAAAGTGGTATCAATGTTATAAAAAAGACAGACAAGTTATGATGCAACAATGTTTTAG	1535
Db	73109	AAAATAAAAAACCCAAAAAAAATAAAATTCAAAATTTTATTAACACTTCAAT	73050
Qy	1536	GAGATTACAATCCATCGGGAAGATGTCAAACTCATCTCTATATATATATATATTTCAAT	1595
Db	73049	AAAATGACAATAAAATCGTTAATTAAGTTTTCACAACAAATTTGATATAAAAAATAATAGTTAAAT	72990

VERSION	X17375.2	GI:7711033
KEYWORDS	16S ribosomal RNA; 16S rRNA gene; 23S ribosomal RNA; 23S rRNA gene; atp6 gene; atp9 gene; ATPase subunit 6; ATPase subunit 9; cob gene; cox1 gene; cox2 gene; cox3 gene; cytochrome b; cytochrome oxidase subunit 1; cytochrome oxidase subunit 2; cytochrome oxidase subunit 3; nad1 gene; nad2 gene; nad3 gene; nad4 gene; nad4L gene; nad5 gene; nad6 gene; NADH dehydrogenase subunit 1; NADH dehydrogenase subunit 2; NADH dehydrogenase subunit 3; NADH dehydrogenase subunit 4; NADH dehydrogenase subunit 4L; NADH dehydrogenase subunit 5; NADH dehydrogenase subunit 6; orf130; ORF148; ORF151; ORF367; ORF390; orf76; ORF90; rnl1 gene; rnl2 gene; rnl3 gene; rnl4 gene; rnl5 gene; rnf2 gene; Scenedesmus obliquus.	
SOURCE ORGANISM	Scenedesmus obliquus Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae; Chlorococcales; Scenedesmeaceae; Scenedesmus.	
REFERENCE AUTHORS	Kuck,U., Godehardt,I. and Schmidt,U.	
TITLE	A self-splicing group II intron in the mitochondrial large subunit rRNA (LSUrRNA) gene of the eukaryotic alga Scenedesmus obliquus	
JOURNAL MEDLINE	Nucleic Acids Res. 18 (9), 2691-2697 (1990)	
REFERENCE MEDLINE	90251444	
AUTHORS	2 (bases 1 to 42781)	
TITLE	Winkler,M. and Kuck,U.	
JOURNAL MEDLINE	The group IIB intron from the green alga Scenedesmus obliquus	
REFERENCE MEDLINE	Curr. Genet. 20 (6), 495-502 (1991)	
AUTHORS	92145774	
TITLE	3 (bases 1 to 42781)	
JOURNAL MEDLINE	Kuck,U., Jekosch,K. and Holzamer,P.	
REFERENCE MEDLINE	DNA sequence analysis of the complete mitochondrial genome of the green alga scenedesmus obliquus: evidence for UAG being a leucine and UCA being a non-sense codon	
AUTHORS	Gene 253 (1), 13-18 (2000)	
TITLE	20384610	
JOURNAL MEDLINE	4 (bases 1 to 42781)	
REFERENCE MEDLINE	Kueck,U.	
AUTHORS	Direct Submission	
TITLE	Submitted (20-DEC-1989) Kueck U., Lehrstuhl fuer Allgemeine Botanik, Ruhr-Universitaet Bochum, Postfach 10 21 40, Universitaetsstr.150, D-44780 Bochum	
JOURNAL MEDLINE	Revised by [4]	
REFERENCE MEDLINE	5 (bases 1 to 42781)	
AUTHORS	Jekosch,K.	
TITLE	Direct Submission	
JOURNAL MEDLINE	Submitted (05-MAY-2000) Jekosch K., Ruhr-Universitaet Bochum, LS fuer Allgemeine und Molekulare Botanik, Universitaetsstr. 150, D-44780 Bochum, GERMANY	
REFERENCE MEDLINE	On May 23, 2000 this sequence version replaced gi:7573226	
AUTHORS	gi:7160201 gi:7619793 gi:7573572 gi:13633.	
TITLE	Location/Qualifiers	
FEATURES	1. .42781 /organism="Scenedesmus obliquus" /organelle="mitochondrion" /strain="KS3-2" /db_xref="taxon:3088" /country="Germany" complement(770..1000) /gene="orf76" complement(770..1000) /gene="orf76" /function="unknown" /codon_start=1 /transl_table=22 /product="ORF76" /protein_id="CAB90358.1" /db_xref="GI:7711034" /translation="MCISTQIKNFYFLSFIFLFIKPFYFLSFIFLFIKPFYFLSFIFLFI KNUYLNLLKKCKADYKSDSFCCFYINIKSLL" 1322..1394 /gene="tRNA-Ile" /product="transfer RNA Ile"	
gene		
CDS		
tRNA		

```
gene
1322..1394
/gene="trna-ile"
/gene="trna-ile"
complement(3340..3732)
/gene="orf130"
CDS
complement(3340..3732)
/gene="orf130"
/function="unknown"
/codon_start=1
/transl_table=22
/product="orf130"
/protein_id="CAB90359.1"
/db_xref="GI:7711035"
/translation="MQOKTLCISLEQGFASGRMGVLDLSRFLGLRLTLVLGP
KGRSTLRLNQRFHFRYLDKGAALLVTRFCSIEQKRLKT'KRAQHFCSLNKRVIG
VFLKIPCKAILLVSRQDQELQVKS"
4893..5342
/gene="cox2"
4893..5342
/gene="cox2"
/function="respiratory chain"
/codon_start=1
/transl_table=22
/product="cytochrome oxidase subunit 2"
/protein_id="CAB90360.1"
/db_xref="GI:7711036"
/translation="MISMVNMCMNRADAPMAWQKLFQDPATSNMGEIVDLHADICFF
LIVLVLMLGARILVSPHHLPVPERENHHTSLLELWAILPSVITVLTIALPSLSL
VNYDDLVSXPALTVKTRQWYWSYAMNEHVQMNLSQQAADLLQS"
join(6092..6802,7414..8655)
/gene="nad5"
join(6092..6802,7414..8655)
/gene="nad5"
join(6092..6802,7414..8655)
/gene="nad5"
/function="respiratory chain"
/codon_start=1
/transl_table=22
/product="NADH dehydrogenase subunit 5"
/protein_id="CAB90377.1"
/db_xref="GI:7711053"
/translation="MYLNAVLPFLGSALAGLGRWLARGSCVTLILNGTSTILSY
VIFETKSMGSAVVLGSGWFGAHTVTEWLLCFDPLSAAAMVTVCTVSCVHIYSLG
YMQNDPLHPLFSLYSLFTGSMELLTANDLLTLMVGEMIGVCYLLILGFWRHLRSR
TKRAQMLNRYSDTVLLFGLFCWMLGSDTSLTATSTAYADWICFALLGGA
LGRSAOVLHRLVADAMEGXPVSAIHAATLVAGTFLIARTNRIWECVYARTVLL
WVGAVTSLMRTWGLQVNDVKRVIAVSTCSQLCYMVVALSLSHVGLAMVHMTACEK
RLFLGAGVYIHACADVQDWRSGGAHSLPLAWTCLLXGSLSLGWPFLAGIYSKDA
ILBLSWATPGSAAALGHLILMTVACLTSAVSFRVLIAVYAPNARKSEVRTPGVPLT
MAYPLCLIALGSIFRYLLSDALIGWTDFWANSIVRSPATTOAVRSHMIPWISALP
LATVAGALLAYVFWVWPLPYCAESYWKPKYLFLOARWGFDLVWNOQISMKVLGAGSAS
WASLDKGLLELLGPGISOKVRGTLPLSPVQKMQTGAVHDYALLLIQILVVVGLLLSY
VAGLTENLLVDYKTLRLVLLISICL"
6092..6802
/gene="nad5"
/number=1
6803..7413
/notes="group I intron"
/number=1
6853..7125
/gene="orf90"
6853..7125
/gene="orf90"
/function="unknown"
/codon_start=1
/transl_table=22
/product="ORF90"
/protein_id="CAB90361.1"
/db_xref="GI:7711037"
/translation="MKFLGGVRLRFSGKGAALWRLRQRFERHFKKGLDPPFLKEAL
FVFDKKEIKELICLVLPHTPKTTLNDLHT'KQNNHNEKFIFIFYIE"
7414..8655
/gene="nad5"
/number=2
```

```
gene
9149..9451
/gene="nad4L"
9149..9451
/gene="nad4L"
/function="respiratory chain"
/codon_start=1
/transl_table=22
/product="NADH dehydrogenase subunit 4L"
/protein_id="CAB90362.1"
/db_xref="GI:7711038"
/translation="MNSALFLCTVVVLFMGALGTFLNRRNFMILICIEVMLLSVNL
VFUTGSVYVDLNGQLLALWVLTAAAEATGALCALCVLYFLRSTLDVELISLMKG"
9858..9940
/gene="trna-tyr"
/notes="codon recognized: UAC"
/product="trna-tyr"
/anticodon=(pos:9892..9894,aa:Tyr)
9858..9940
/gene="trna-tyr"
10046..10127
/gene="trna-leu"
/notes="codon recognized: CUU"
/product="trna-leu"
/anticodon=(pos:10079..10081,aa:Leu)
10046..10127
/gene="trna-leu"
10133..10203
/gene="trna-arg"
/notes="codon recognized: AGG"
/product="trna-arg"
/anticodon=(pos:10165..10167,aa:Arg)
10133..10203
/gene="trna-arg"
10213..10284
/gene="trna-ile"
/notes="codon recognized: AUC"
/product="trna-ile"
/anticodon=(pos:10245..10247,aa:Ile)
10213..10284
/gene="trna-ile"
10486..10839
/gene="nad3"
/function="respiratory chain"
/codon_start=1
/transl_table=22
/product="NADH dehydrogenase subunit 3"
/protein_id="CAB90363.1"
/db_xref="GI:7711039"
/translation="MTEYVGVLLYIAVGLGIVILFRSYKVRPRRLDEKATAYECG
```

```
Query Match 3.0%; Score 49; DB 14; Length 42781;
Best Local Similarity 58.6%; Pred. No. 0.17;
Matches 85; Conservative 0; Mismatches 60; Indels 0; Gaps 0;

Qy 1476 AAAAACTGCTATCAATGTATAAAAAGACAGACAAGTATGATGACACAAATCTTTTAA 1535
Db 22996 AAAAAATTTTTTATTTGTATAAAGATACATACATAAAAAATAAAATGTTTTATTG 23055

Qy 1536 GAGATTACAAATCCATGGGAAGATGTATCAACATCATCTCTCTATATATATATATCAAT 1595
Db 23056 TATAAAGATATACACATATATAAAATAAAATGTTACTTTAAATATATATATATATA 23115

Qy 1596 TGTTTTAAAAAATAAAAAA 1620
Db 23116 TATATTTTAAAAAATAAAAAA 23140
```

```
RESULT 13
AF204057 42919 bp DNA circular PLN 22-JUN-2000
LOCUS Scenedesmus obliquus strain UTEx 78 mitochondrion, complete genome.
DEFINITION
```

```

ACCESSION AF204057
VERSION AF204057.1 GI:8099190
KEYWORDS
SOURCE Scenedesmus obliquus.
ORGANISM Mitochondrion Scenedesmus obliquus
Eukaryota; Viridiplantae; Chlorophyta; Chlorophyceae;
Chlorococcales; Scenedesmaceae; Scenedesmus.
REFERENCE 1 (bases 1 to 42919)
AUTHORS Nedelcu, A.M., Lee, R.W., Lemieux, C., Gray, M.W. and Burger, G.
TITLE The complete mitochondrial DNA sequence of scenedesmus obliquus
reflects an intermediate stage in the evolution of the green algal
mitochondrial genome
JOURNAL Genome Res. 10 (6), 819-831 (2000)
MEDLINE 20314873
REFERENCE 2 (bases 1 to 42919)
AUTHORS Burger, G.
TITLE Direct Submission
JOURNAL Submitted (11-NOV-1999) Departement de Biochimie, Universite de
Montreal, 2900 Boulevard Edouard-Montpetit, Montreal, Quebec H3T
1J4, Canada
FEATURES
source Location/Qualifiers
1..42919
/organism="Scenedesmus obliquus"
/organelle="mitochondrion"
/strain="UTEX 78"
/db_xref="taxon:3088"
1..716
/rRNA /gene="rnl_a"
/note="similar to Escherichia coli sequence"
/product="large subunit ribosomal RNA"
1..716
/gene="rnl_a"
786..858
/rRNA /gene="trnA(ugc)"
/product="tRNA-Ala"
/anticodon="(pos:819..821,aa:Ala)
786..858
/gene="rnl_a"
Join(859..1871,2480..2528,3802..4158)
/rRNA /gene="rnl_d"
/note="similar to Escherichia coli sequence"
/product="large subunit ribosomal RNA"
859..4158
/gene="rnl_d"
859..1871
/exon /gene="rnl_d"
/number=1
/intron /gene="rnl_d"
1872..2479
/exon /gene="rnl_d"
/note="Group IIB; not present in Escherichia coli
sequence"
/number=1
/exon 2480..2528
/gene="rnl_d"
/number=2
/intron 2529..3801
/gene="rnl_d"
/note="Group IAL; not present in Escherichia coli
sequence"
/number=2
/gene 3027..3674
/gene="orf215"
3027..3674
/CDS /gene="orf215"
/gene="orf215"
/note="Orf inside an intron"
/codon_start=1
/transl_table=22
/product="orf215"
/db_xref="GI:8099191"
/db_xref="GI:8099191"
/transl_table=22
/translation="MLPPFLFEVAIGNRIGDRSLYTTTKGSKIKLEQGRVKNKDYLYH
LPFLYKGMTNIEKPYRIPKVTGTYTRGVKISYFRITTHPAFDKIYNFFTCNGKKT
YREGLIITNHLTSVGLSYWVRDQGLQKNNEIILRTMGFTQKENIQIIVEELNRKFKLNS
KVKCHKKYWAIYIPASDAPTIRKHLVLPESIKYKMKPLIKKHRSLYLYIFIH"
3802..4158
/exon /gene="rnl_d"
/number=3
/rRNA /gene="trnH(gug)"
/product="tRNA-His"
/anticodon="(pos:4368..4370,aa:His)
4335..4407
/gene="rnl_c"
4335..4407
/rRNA /gene="trnH(gug)"
/note="similar to Escherichia coli sequence"
/product="large subunit ribosomal RNA"
4644..5018
/gene="rnl_c"
4644..5018
/rRNA /gene="trnN(guu)"
/product="tRNA-Asn"
/anticodon="(pos:5117..5119,aa:Asn)
5084..5156
/gene="rnl_c"
5084..5156
/rRNA /gene="trnM(cau)"
/product="tRNA-Lys"
/anticodon="(pos:6099..6101,aa:Lys)
5396..5469
/gene="rnl_c"
5396..5469
/rRNA /gene="trnK(acy)"
/product="tRNA-Arg"
/anticodon="(pos:5430..5432,aa:Arg)
5396..5596
/gene="rnl_c"
5396..5596
/rRNA /gene="trnM(cau)"
/product="tRNA-Met"
/anticodon="(pos:5557..5559,aa:Met)
5525..5596
/gene="rnl_c"
5525..5596
/rRNA /gene="trnK(uuu)"
/product="tRNA-Lys"
/anticodon="(pos:6099..6101,aa:Lys)
6066..6138
/gene="rnl_c"
6066..6138
/rRNA /gene="trnK(uuu)"
/product="tRNA-Lys"
/anticodon="(pos:6099..6101,aa:Lys)
6448..7107
/gene="nad6"
6448..7107
/CDS /gene="nad6"
/EC_number="1.6.5.3"
/codon_start=1
/transl_table=22
/product="NADH dehydrogenase subunit 6"
/protein_id="AAF72047.1"
/db_xref="GI:8099192"
/transl_table=22
/translation="MSSSFALITATGALMSVFLMMSKNPVHSVFYLVLLFLHCSGLLY
LLGLEFYVLQLLYVYGALAILFLVWMLDIPSTEILRHQGTTPVAGVGLGLCLLS
ALLAFTQPLVESFPQTSPIPTSTFGQNSDFVSSRSRYSFNWNTRTATSTVANLGR
LYGVHADLIIRSLLLVAMVAVGLTLKRPVSPNQDVFRQHYVDFQKSLVQIPQNK
I"
7148..8347
/gene="cob"
7148..8347
/CDS /gene="cob"
/EC_number="1.10.2.2"
/codon_start=1
/transl_table=22
/product="apocytochrome b"
/protein_id="AAF72048.1"
/db_xref="GI:8099193"
/transl_table=22
/translation="MKTSLKLRGSLWDOPMLAVAOHNLISYPTPSNLNGYNNMGVLA
GLCLVLIQTGIFLAMYTAHVDLAFHSYQHLMDRVPNGWLLRYLHANGRSLEFFIVVY
MHLFRGLYTSVAOPREFVWLVGVVILLIMLTAFYVLPWGQSLMAGRTVITSLAS
ALPLVGTIVGWLWGGFAVDNPTLNFYSFHYLFEPFILLRSLHLRALHGYSTNPL
GNAQSDLVTFYPYPTKDLVAALRLSFGAAYLVNFPPELLGHDPNNIPANPYATPRH
IYPENYFLWVYAILRSIPNKLAVRAIGLVFAALAAPTMHHPKPCRGSPFRKLHVLV
WALIGDFMLLSYLGQOPVEPPYVVLGQRTTVEFFSLILGLCVRWADALLANFSLKKS
HSKDFAC"

```

[illegible]

DEFINITION	Plasmodium falciparum MAL4P2, complete sequence.
ACCESSION	AL035475
VERSION	AL035475.6 GI:6562702
KEYWORDS	HTG: DNA polymerase alpha; DnaJ; flap exonuclease; protein phosphatase; replication A protein; ribosomal processing protein; ribosomal protein; rifin; transcription factor; var. malaria parasite P. falciparum.
SOURCE	Plasmodium falciparum
ORGANISM	Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
REFERENCE	1 (bases 1 to 234112)
AUTHORS	Devlin,K., Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M. and Barrrell,B.
TITLE	Direct Submission
JOURNAL	Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
COMMENT	On Dec 12, 1999 this sequence version replaced gi:5731898. For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/projects/P.falciparum. IMPORTANT: This sequence is not the entire insert of clone MAL4P2. It may be shorter because we only sequence overlapping sections once, or longer because we arrange for a small overlap between neighbouring submissions.
FEATURES	Location/Qualifiers
Source	1..234112 /organism="plasmodium falciparum" /strain="3D7" /db_xref="taxon:5833" /chromosome="4" /clone="MAL4P2" /complement(1..3390) /gene="MAL4P2.01" complement(1..3390) /partial /gene="MAL4P2.01" note="MAL4P2.01, hypothetical protein, partial orf, len: > 1130 aa, contains NKSPPD repeat sequence"
gene	/codon_start=1 /product="hypothetical protein MAL4P2.01" /protein_id="CAB62842.1" /db_xref="GI:6562703" /db_xref="SPTREMBL:Q9U0M1"
CDS	/translation="MDQKNLEICGDNSKEETKEHNVLLGVKTKNHLNEEEKLVDD KTYNPKSLNIINIANKNDIMLTQVKNDKTLKDDTIERINOKVVTKNHLNEEKK THEDINQNGNKSPPDNKSKPDNNKSKPDNNKSKPDNNKSKPDNNKSKPDNNKSKPDNNK SKPDNNKSKPDNNKSKPDNNKSKPDNNKSKPDNNKSKPDNNKSKPDNNKSKPDNNKSPNDNK SNPDNNKSPNDNKSPNDNKSPNDNKSPNDNKSPNDNKSPNDNKSPNDNKSPNDNKSPNDNK ENIRKQHFMKYKYNITYHAEKYMNIEIKRYEENISFARSYLCHDLNLINKDT INFVKRRLNELPVDIRNLCSEHNKNDKIFENEINGLSILSEDKTNNMEIKDDEYYIKIE QTNHOCFCEDNVNPVDIWNIIHKDTVNATNINKEDDNYLDMDHIHMKEWNSVGL SVIYYMKNDNMESFNKVYKIYIMKSGENVTRHMKDIDFORCIYNTCSKNITYSIHV SDNILCLGNVKGDETFVNPYPYKKNWEIKSTNNYNNINRDKNRTHSLKANVR LIINHKNISYVICNKLAKGKNYERRMLKNMLKYNLIKINLYILSCNNIYKYAKK KILOSQNFNIKTNINDMYNKINTIIGKEKEKDNNIKHTKHNSKHNNNNNN NKPSPNDIOINVMKKRTYIPFFINTYKLFTHYKNI SFKEFLMRKFVAKYDIYIF NKLZYITLKIEDLCKLYMDNALEILDNLTDHIKENQLONLCHTVLINREIENI KQPLHFLEYSKTHSRFLSKHEWTNKYIFLNQKKENTQNDHISPLEVINQSIE EQMDLHYWTHYTHYIYEKSMFAYGKIAQRNLEIMWKCRISBQINELEKKIK REIMTEBOEKMAKKEKPIFEBKXKIDQKIYIDEENKMMTERELIIDQLINKHDT KNNTINKEMLEKEIKNFSEKNTITEKQNOSDKITTEKEIQNIRNEKMTLEKIEI INSNDKITTEKEIQNFRDKITTEKEIKNFRNDKMTLEKEIKNFSNDKITTEKEIQN NREKITTEKEIQNISNDKITTEKEIQNIRDNTLEKEIQNF"
misc_feature	complement(3444..3450) /note="(G7)--tract upstream of MAL4P2.01"
gene	5644..6063 /gene="MAL4P2.02"
CDS	5644..6063 /gene="MAL4P2.02" note="MAL4P2.02, very hypothetical protein, len: 139 aa" /codon_start=1 /product="very hypothetical protein, MAL4P2.02" /protein_id="CAB62843.1" /db_xref="GI:6562704"


```

LOCUS       AC023268      144457 bp      DNA      HTG      01-MAR-2000
DEFINITION  Homo sapiens chromosome 9 clone RP11-168D21 map 9, WORKING DRAFT
SEQUENCE, 22 unordered pieces.
ACCESSION   AC023268
VERSION     AC023268.2  GI:7139920
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      human.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 144457)
Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beckerly,R., Bieda,F.,
Boguslavsky,L., Boukhgaiter,B., Brown,A., Burkett,G., Castile,A.,
Choepel,Y., Colangelo,M., Collins,S., Collumore,A., Cooke,P.,
Dearellano,K., Dewar,K., Domino,M., Doyle,M., Fenesfor,J.,
Ferrelira,P., FitzHugh,W., Forrest,C., Gage,D., Galagan,J.,
Gardyna,S., Grant,G., Hagos,B., Heaford,A., Horton,L.,
Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
Landers,T., Lehoczy,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
Macdonald,P., Marquis,N., McEwan,P., McGurk,A., McKernan,K.,
McPheeters,R., Meldrim,J., Meneus,L., Morrow,J., Naylor,J.,
Norman,C.H., O'Connor,T., O'Donnell,P., Olivari,T.M., Peterson,K.,
Pierre,N., Pisani,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thomann,N.,
Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye.W.J.,
Zimmer,A. and Zody,M.
Direct Submission
Submitted (10-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Mar 1, 2000 this sequence version replaced gi:6938035.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
http://ftp.genome.washington.edu/RM/RepeatMasker.html
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WtBR
Web site: http://www.seq.wi.mit.edu
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L879
Center clone name: 168_D_21
----- Summary Statistics
Sequencing vector: M13; M77815; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 133287 bases at least Q40
Consensus quality: 139023 bases at least Q30
Consensus quality: 141009 bases at least Q20
Insert size: 163000; agarose-fp
Quality coverage: 3.9 in Q20 bases; sum-of-contigs
Quality coverage: 4.5 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
1
* 2174 2273: contig of 2173 bp in length
* 2274 4319: contig of 100 bp
* 4320 4419: contig of 2046 bp in length
* 4420 5565: contig of 100 bp
* 5566 6665: contig of 2146 bp in length
* 6666 9518: contig of 100 bp
* 9519 9618: contig of 100 bp
* 9619 13288: contig of 100 bp
* 13289 13388: contig of 100 bp
* 13389 15599: contig of 2211 bp in length
* 15600 15699: gap of 100 bp
* 15700 21304: contig of 5605 bp in length
* 21305 21404: gap of 100 bp
* 21405 26070: contig of 4666 bp in length
* 26071 26170: gap of 100 bp
* 26171 31773: contig of 5603 bp in length
* 31774 31873: gap of 100 bp
* 31874 38024: contig of 6151 bp in length
* 38025 38124: gap of 100 bp
* 38125 43487: contig of 5363 bp in length
* 43488 43587: gap of 100 bp
* 43588 49049: contig of 5462 bp in length
* 49050 49149: gap of 100 bp
* 49150 56108: contig of 6959 bp in length
* 56109 56208: gap of 100 bp
* 56209 64181: contig of 7973 bp in length
* 64182 64281: gap of 100 bp
* 64282 73216: contig of 8935 bp in length
* 73217 73316: gap of 100 bp
* 73317 81917: contig of 8601 bp in length
* 81918 82017: gap of 100 bp
* 82018 91505: contig of 9488 bp in length
* 91506 91605: gap of 100 bp
* 91606 101191: contig of 9586 bp in length
* 101192 101291: gap of 100 bp
* 101292 110389: contig of 9098 bp in length
* 110390 110489: gap of 100 bp
* 110490 120867: contig of 10378 bp in length
* 120868 120967: gap of 100 bp
* 120968 131985: contig of 11018 bp in length
* 131986 132085: gap of 100 bp
* 132086 144457: contig of 12372 bp in length.
FEATURES             Location/Qualifiers
     source            1..144457
                        /organism="Homo sapiens"
                        /db_xref="taxon:9606"
                        /map="9"
                        /clone_lib="RP11-168D21"
                        /clone="RP11-168D21"
     misc_feature      1..2173
                        /note="assembly_fragment"
     misc_feature      2274..4319
                        /note="assembly_fragment"
     misc_feature      4420..6565
                        /note="assembly_fragment"
     misc_feature      6666..9518
                        /note="assembly_fragment"
     misc_feature      9619..13288
                        /note="assembly_fragment"
     misc_feature      13389..15599
                        /note="assembly_fragment"
                        clone_end:SP6
                        vector_side:right
     misc_feature      15700..21304
                        /note="assembly_fragment"
     misc_feature      21405..26070
                        /note="assembly_fragment"
     misc_feature      26171..31773
                        /note="assembly_fragment"
     misc_feature      31874..38024
                        /note="assembly_fragment"
     misc_feature      38125..43487
                        /note="assembly_fragment"
     misc_feature      43588..49049
                        /note="assembly_fragment"
     misc_feature      49150..56108
                        /note="assembly_fragment"

```

```
misc_feature 56209..64181
/note="assembly_fragment"
misc_feature 64282..73216
/note="assembly_fragment"
misc_feature 73317..81917
/note="assembly_fragment"
misc_feature 82018..91505
/note="assembly_fragment"
misc_feature 91606..101191
/note="assembly_fragment"
misc_feature 101292..110389
/note="assembly_fragment"
misc_feature 110490..120867
/note="assembly_fragment"
misc_feature 120968..131985
/note="assembly_fragment"
misc_feature 132086..144457
/note="assembly_fragment
clone_end:T7
vector_side:right"
BASE COUNT 41302 a 30843 c 30435 g 39775 t 2102 others
ORIGIN
```

```
Query Match 3.0%; Score 48.2; DB 67; Length 144457;
Best Local Similarity 57.7%; Pred. No. 0.33;
Matches 86; Conservative 0; Mismatches 63; Indels 0; Gaps 0;

QY 1474 AGAAAAAGTGTATCAATGTATATAAAGACAGACAGAGTATGATGCGAACAAATCTTTT 1533
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111287 AGTGAACCTTTGTCTCAGAAAAAAGAAAAAGAAAAAGAAAAAGAAATATCACTTTT 111228

QY 1534 AGGAGATTACAATCCATGGGAGATGTATCAAACTCATCTCTATATATATATATATCA 1593
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111227 AGAGCTTAGGATTAGGCTATATGTGACGACACTAAATTTCTTTAAAAATATCTTTA 111168

QY 1594 ATGTTTTTAAAAAAGAAAAAAGAAAAA 1622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 111167 TTTTATTAAACATTACAATTGGAAAAA 111139
```

Search completed: November 5, 2001, 16:38:25
Job time: 10075 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 13:52:40 ; Search time 5816.79 seconds
(without alignments)
2767.542 Million cell updates/sec

Title: US-08-894-356C-1
Perfect score: 1703
Sequence: 1 TCATTATGGAGCAATCCAA.....AAAAAAAAAAAAAAAAAAAA 1703

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:
2:	gb_est2:
3:	gb_est3:
4:	gb_est4:
5:	gb_est5:
6:	gb_est6:
7:	gb_est7:
8:	gb_est8:
9:	gb_est9:
10:	gb_est10:
11:	gb_est11:
12:	gb_est12:
13:	gb_est13:
14:	gb_est14:
15:	gb_est15:
16:	gb_est16:
17:	gb_est17:
18:	gb_est18:
19:	gb_est19:
20:	gb_est20:
21:	gb_est21:
22:	gb_est22:
23:	gb_est23:
24:	gb_est24:
25:	gb_est25:
26:	gb_est26:
27:	gb_est27:
28:	gb_est28:
29:	gb_est29:
30:	gb_est30:
31:	gb_est31:
32:	gb_est32:
33:	gb_est33:
34:	gb_est34:
35:	gb_est35:
36:	gb_est36:
37:	gb_est37:
38:	gb_est38:
39:	gb_est39:
40:	gb_est40:
41:	gb_est41:
42:	gb_est42:
43:	gb_est43:

44:	em_esthum10:
45:	em_esthum11:
46:	em_esthum12:
47:	em_esthum13:
48:	em_esthum14:
49:	em_esthum15:
50:	em_esthum16:
51:	em_esthum17:
52:	em_esthum18:
53:	em_esthum19:
54:	em_esthum20:
55:	em_esthum21:
56:	em_esthum22:
57:	em_esthum23:
58:	em_esthum24:
59:	em_esthum25:
60:	em_esthum26:
61:	em_esthum27:
62:	em_esthum28:
63:	em_estin1:
64:	em_estin2:
65:	em_estin3:
66:	em_estin4:
67:	em_estin5:
68:	em_estom1:
69:	em_estom2:
70:	em_estov1:
71:	em_estov2:
72:	em_estpl1:
73:	em_estpl2:
74:	em_estpl3:
75:	em_estpl4:
76:	em_estpl5:
77:	em_estpl6:
78:	em_estpl7:
79:	em_estpl8:
80:	em_estpl9:
81:	em_estpl10:
82:	em_estro1:
83:	em_estro2:
84:	em_estro3:
85:	em_estro4:
86:	em_estro5:
87:	em_estro6:
88:	em_estro7:
89:	em_estro8:
90:	em_estro9:
91:	em_estro10:
92:	em_estro11:
93:	em_estro12:
94:	em_estro13:
95:	em_estro14:
96:	em_estro15:
97:	em_estro16:
98:	em_estro17:
99:	em_estro18:
100:	em_estro19:
101:	em_estro20:
102:	gb_est25:
103:	gb_est26:
104:	gb_est27:
105:	gb_est28:
106:	gb_est29:
107:	gb_est30:
108:	gb_est31:
109:	gb_est32:
110:	gb_est33:
111:	gb_est34:
112:	gb_est35:
113:	gb_est36:
114:	gb_est37:
115:	gb_est38:
116:	gb_est39:

```
117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*
```

```
190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

TITLE Generation of ESTs from Potato Leaves Challenged with Phytophthora
infestans, Incompatible Reaction
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES
source

1..763
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="BPL110A14"
/clone_lib="P. infestans-challenged leaf"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"

BASE COUNT 212 a 162 c 156 g 233 t
ORIGIN
Query Match 8.3%; Score 141.2; DB 155; Length 763;
Best Local Similarity 52.6%; Pred. No. 8.e-27;
Matches 382; Conservative 0; Mismatches 333; Indels 11; Gaps 3;
QY 78 GAGTTATCGTACCGGTAAACATCTTTCGATATCCCTGGTTCGACATTTGAATAGATGCAG 137
DB 35 GAGGTACACTCCCTCTTACTTATTGATCATGTTGGTAGGGTCCACCGTATACGG 94
QY 138 TCCTTCTGTTTACGACTTTCGGTACCAAGAACACATTTCTTGGACACTGTTATCCCT 197
DB 95 CGGATATTATCTACAAAGCTCCCACTTTCCAAACCGGATTTCTGTTCAAAACATATTCT 154
QY 198 ATCTTTAAGGCCTCTTGTCTCTCACTTAAACACACTAGCTTCGGCTAGCGGAATTTG 257
DB 155 ACTCTTTAAATAATTCACCTCTCCCTCACTCTCAACACACTACACGCCCTTAGCAGGAATGTT 214
QY 258 TTGATCGCGATCAAAATCGGCGAAATGCCGAGTCTTCAGTACTCCGCTGATAGGGCGAC 317
DB 215 GCTCGTCCACTAGATACAATGGATATCTGAGTTACGTTA-----TGTGACAGGAAT 268
QY 318 TCGATAACTTTGATCGTTGGGAGTCTGACAGGATTTTGTACCTTAAAGTCTATCA 377
DB 269 TCTGTATCTGTTACTTTTTTGGAGACTGATATGGATTCAATTTCAATTTATCTCATTTGTCACCAT 328
QY 378 CTGGTAGATTTCAATGATTTGATGGCTTTTATGTTATGTCACCGGTTTATAAGGACC 437
DB 329 CCAGGTAATGCTAAGGATTTTAT---CACTTTGTTCTTAAGTTAGGGGAACCTTAAGGAT 385
QY 438 ATGCAAGACTATAAAGTGTATCCCGCTCGTAGCGTCAAGTAAACCGTTTTTCTTAACCGT 497
DB 386 GCACCGGGGTCCAATAGTCCCGCTCTAGCCATTTCAAGTGACACTTTTCCGCAATCTT 445
QY 498 GGATAGCCGTGGCTCTGACGGGACATCATTTCAATTTGCAGATGCTAAAAGTTTTTGTAAATG 557
DB 446 GGTGTATCCATTTGTTTCACTAACCACTCATGTTGTTGGTGTGAGCTACCATAGTAGGG 505
QY 558 TTCATCAATGCTTTGGGCTTATATTACAATAATTGGGAAGACACCGGACTTTGTTGCCGG 617
DB 506 TTCAATTAAGGCGTGGGCTCTACTCCATAAATTCGGTGGAGATGAACAATTTCTTATCGAAT 565

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	141.2	8.3	763	155	BG591758	EST499600
2	141	8.3	575	142	BE920468	EST424237
3	127	7.5	690	155	BG599447	EST504342
4	122.8	7.2	591	167	BE434257	EST405335
5	121	7.1	574	142	BE922746	EST426515
6	119.6	7.0	568	142	BE921494	EST425179
7	111.4	6.5	504	173	BG097054	EST461573
8	111.4	6.5	578	143	BF051101	EST436276
9	109.8	6.4	606	113	AW221049	EST297518
10	107.8	6.3	570	142	BE922784	EST426553
11	107.6	6.3	603	118	AW616119	EST296884
12	107.2	6.3	542	167	BE436185	EST407263
13	102.8	6.0	659	155	BG598154	EST496832
14	101.6	6.0	497	122	AW930934	EST356777
15	101	5.9	583	113	AW221050	EST297519
16	97.4	5.7	549	21	AT490445	EST248771
17	93.6	5.5	687	111	AW102336	sd86d06.y
18	91.6	5.4	404	120	AW737239	EST338666
19	91.6	5.4	543	119	AW650650	EST329104
20	90.8	5.3	518	119	AW650280	EST328734
21	89.4	5.2	783	155	BG584708	EST486469
22	87.4	5.1	411	142	BE922378	EST426147
23	86.6	5.1	414	142	BE921952	EST425721
24	85.4	5.0	526	105	AL386584	MTBC35E10
25	85.2	5.0	705	114	AW299055	EST305729
26	84.6	5.0	788	164	BE187619	EST336180
27	84	4.9	411	142	BE921804	EST425573
28	82.8	4.9	487	118	AW616206	EST307245
29	81.4	4.8	368	142	BE921953	EST425722
30	80.2	4.7	403	143	BF053202	EST438432
31	78.4	4.6	648	154	BG457806	NF034C06P
32	76.6	4.5	386	173	BG097680	EST462199
33	75.8	4.5	451	167	BE433920	EST404998
34	75.6	4.4	505	154	BG508632	sac75409
35	75.2	4.4	422	155	BG598694	EST497536
36	75	4.4	708	138	BE660849	7-C8 Gmax
37	74.6	4.4	605	30	AV440958	AV440958
38	74.2	4.4	550	105	AL380276	MTBB51E01
39	74	4.3	482	144	BF096979	EST400516
40	73.8	4.3	468	167	BE461227	EST412646
41	73	4.3	403	143	BF052576	EST437806
42	70.2	4.1	485	105	AL386583	MTBC35E10
43	69.8	4.1	408	142	BE920022	EST423791
44	69.2	4.1	569	119	AW719572	LJNEST6h1
45	68	4.0	561	121	AW831571	sm34h07.y

ALIGNMENTS

RESULT 1	BG591758	BG591758	763 bp	mRNA	EST	12-APR-2001
LOCUS	EST499600	P. infestans-challenged leaf	Solanum tuberosum	cdna	clone	
DEFINITION	BPL110A14	5' sequence, mRNA sequence.				
ACCESSION	BG591758					
VERSION	BG591758.1	GI:13609898				
KEYWORDS	EST.					
SOURCE	potato.					
ORGANISM	Solanum tuberosum					
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;					
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;					
	Asteridae; euasterids I; Solanales; Solanaceae; Solanum.					
	1 (bases 1 to 763)					
	Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukolanov, A.,					
	Rangel, P., Haberlach, G.T., Cho, J., Chienlingo, A., Bougri, O., Buehl					
	, C.R., Ronning, C.M., Helgeson, J. and Baker, B.					

```

Qy 618 AATCTTCTTCATCTTCGATAGATCGAATAATCAAGAATCTGTATGCCCTAGAGGAACA 677
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 566 GAGCTAATTCATTTATGATAGTCCGTAGTAAGGACCCATCGACAAAGGATGTC 625
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 678 TTTGGACGAAGTCAAGATGTTCTTCAATGTTCTCTAGATTTGGNAGCAAAACCCCT 737
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 626 ATGTGGGAAGAAATGAGAAAGAAATCTAAATATGCGTGACATTTATGACTCCTCCCT 685
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 738 CGATTCAACAAGGTACGAGCTACATATGCTCCTCTCCCTTGCTGAAATCCAGAACTAAAG 797
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 686 GAA--CACAAATGTCGAGGTACATTACTATAAAGAGATGATATTGAGAACTCAG 743
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 798 AACAAA 803
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 744 ATTTAA 749
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 2
BE920468 575 bp mRNA EST 02-OCT-2000
LOCUS EST424237 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION CSTB608 5' sequence, mRNA sequence.
ACCESSION BE920468
VERSION BE920468.1 GI:10446544
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 575)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.

FEATURES
source
    1..575
        /organism="Solanum tuberosum"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="cSTB608"
        /clone_lib="potato leaves and petioles"
        /tissue_type="leaflets and petioles"
        /dev_stage="8 weeks old plants"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
xhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 167 a 77 c 139 g 192 t
ORIGIN
    1..575
        /organism="Solanum tuberosum"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="cSTB608"
        /clone_lib="potato leaves and petioles"
        /tissue_type="leaflets and petioles"
        /dev_stage="8 weeks old plants"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
xhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
Query Match 8.3%; Score 141; DB 142; Length 575;
Best Local Similarity 58.4%; Pred. No. 9.2e-27;
Matches 331; Conservative 0; Mismatches 215; Indels 21; Gaps 4;

Qy 833 AATACGTGTACGACGTTCAATGACGTGTGGATAGTATGATGACATGCGTCAAAATC 892
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 6 AACTCATGTAAACATCTTTTACTGTAACGAGTGCCTTATGTATGGACTTGCCTTGATATAATC 65
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 893 AAAAGATGACGCTCGTATCAGAGGATCATCGAACGCAAAATCAGCTCGGACTACTCAG 952
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 66 AAAGGACGCGATTGG---AGAGAAATGATCATAGATGATGTATGATGAGTCTTTGG 122
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 953 TTTTACAGCGGANTGCCGAGGACTTCTGACGCCGCCCGCTGTCGCCCTAACTACTTTGGCAA 1012
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 123 ATGTGCTGGAGATTTTAGAGCAGCATTCATCAATCCACCACCTTCCTCAATCTTATTTGGGAA 182
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1013 CTGTCTTCGCGTCATGCGCTGCAAAAGCAACA--CATAAAGAGTTAGTTGGGATAAAG 1069
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 183 TTGTGTAGTTGGGTGTGTCACAAGATCAATAAGCATGTTGACTTAGTTGGAAGGAAG 242
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1070 GCTTCTTGTGTCAGTTGACGCTATTGAGAGAGCCATTGAAAAGAGGTTGCACACAGAAA 1129
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 243 TTTTGAATTCGCGTGAATTAATTGGAGAGTCAATTCAAAAAATAATGAAGATGAAGA 302
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1130 AGGCGTTCTTTCAGATGCAAAACTTGGTTATCGGAATCTAATGGAATCCCTTCAAAAAG 1189
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 303 ATGGTCTCTTAATGGTG-----ATTGGTTAAAAGAGTTGGACAACGTAGACCTAATTCG 356
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1190 ATTTCTCGGATTACCGGATCGCCTAAGTTCGATTCGATGCTAGATTTTGGATGGG 1249
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 357 ATTTTCTCAATTCGTCGATCGCCAAACATGACTTCTATGCTGCTGATTTTGGATGGG 416
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1250 AAGCCTGCAAAATTTGACATTACCTCTGTTGATTATGACGAA-----TTGATTTA 1300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 417 AAGGCGCTGAAGTTAGAAATTCATTTCTTGACAATGATGATGGTGAATTCGATGTC 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1301 TGTGATTCAAGTCCAGGATTTTGAAGAGGTGTGGAGATTTGGAGTATCATTTGCCCTAAGAT 1360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 477 TCTTAGTAAATCTAAGAAATTTTGATCGAGATTAGAGATTGTTGCTTGTCTATTATAC 536
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 1361 TCATATGATGTCATTTGCCAAAATCTTT 1387
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 TCAAATGAATGCTTTTGTGCTGCTATATT 563
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 3
BG599447 690 bp mRNA EST 12-APR-2001
LOCUS EST504342 cSTS Solanum tuberosum cDNA clone cSTS25122 5' sequence,
DEFINITION mRNA sequence.
ACCESSION BG599447
VERSION BG599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 690)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES
source
    1..690
        /organism="Solanum tuberosum"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="cSTS25122"
        /clone_lib="cSTS"
        /tissue_type="sprouting eyes from tubers"
        /dev_stage="12-14 weeks post harvest"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
xhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN
    1..690
        /organism="Solanum tuberosum"
        /cultivar="Kennebec"
        /db_xref="taxon:4113"
        /clone="cSTS25122"
        /clone_lib="cSTS"
        /tissue_type="sprouting eyes from tubers"
        /dev_stage="12-14 weeks post harvest"
        /lab_host="SOLR"
        /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
xhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

```

Query Match		7.5%;	Score 127;	DB 155;	Length 690;
Best Local Similarity		52.9%;	Pred. No. 5.1e-23;		
Matches 369;		Conservative 0;	Mismatches 315;	Indels 13;	Gaps 4;
QY	40	AATGCCAAGTTACACACCATCTGCACAAACAGATGTCGAGTTATCGCTACCGGTAAACAT	99		
Db	4	AATGTCAGTTGCGCCACCTCCGGGGCGCA--GCGAGGTGACACTCCCTCTTACTT	60		
QY	100	TCTTCGATATCCCTCGTTCGCACTTGAATAAGATGCAGTCCCTCTCTGTTTACGACTTTC	159		
Db	61	ATTTTGATCATGTTGTTAGGTTCCACCGTATACGGGGATATATTCTACAAGCTCC	120		
QY	160	CGTACCCAGAAGACACATTTCTTGGACACTGTTATCCCTAATCTTAAGGCTCTTTGTCTC	219		
Db	121	CCATTTCCAAACCGGATTTTCGTTTCAAAACATTTATCTCTCTTAAAAATTCACCTCTCC	180		
QY	220	TCACCTCTAAACACTACGTTCCGCTTAGCGGAATTTGTTGATGCCGATCAAAATCGGGCG	279		
Db	181	TCACCTCTCAACACTATACGCCCTTAGCTGGAACGTTGCTGTCACCATAGATACAACG	240		
QY	280	AAATGCCGAAGTTTACGTACTCCGCTGATGAGGGGACACTGCGATAACTTTTGATCGTTGCGG	339		
Db	241	GATATCTGAGTTACATTA-----TGTGACAGGAGATCTCTATCTGTTACTTTTCTG	294		
QY	340	AGTCTGACCAGGATTTTGACTACCTTAAGGTCATCAACTGGTAGATTTCCAATGATTTCG	399		
Db	295	AGACTGATATGGATTTTCAATATATCTCATTTGTTGACCATCCGGGTAATGCTAAGGATTTT	354		
QY	400	ATGGCCTTTTATGTTATGCCACGGGTTATAAGGACCATGCAAGACTATAAAGTAGTCC	459		
Db	355	AT---CACTTTGTCTACGTTTAGGGGAACCTTAAGGATGCACCCGGGTCCAATTAGTCTC	411		
QY	460	CGCTCGTAGCGTGCAAGTAACGGTTTTTCCCTAAACGGTGGCATAGCCGTGGCTCTGACGG	519		
Db	412	CGGTCTTAGCCATTCAAGTGACACTTTTTCCGAATCTTGGTGATCCATTGGTTTCACTA	471		
QY	520	CACATCATTCATTCGAGATGCTTAAAGTTTTTGTATGTTTCATCATGCTTGGGCTATA	579		
Db	472	ACCATCATGTTGTTGGTATGGAGCTACCATAGTAGGGTTTCAATTAGGGCGTGGGCTCTAC	531		
QY	580	TTAACAAATTTGGGAAGACGGGACTTGTGTCGCGAATCTTCTTCGACTTTTCGATA	639		
Db	532	TCCATAAATTCGGTGAGATGAACAATTTCTATCGAATCAGCTAATTCATTTTATGATA	591		
QY	640	-GATCGATAATCAAGATCTGTATGCCCTAGAGGAACAATTTTGGAAACGAATGCAAGAT	698		
Db	592	GGTCCGTAGTAAAGACCCATATGACAAGGATGTTTCATATGGGAAGAAATGAAGAA	651		
QY	699	GTTCTTGAATGTTCTCTAGATTGGGAACAAACCCC	735		
Db	652	AGAATCAAGATATGCGTGACATATTGACTCTCTCTCC	688		
RESULT	4				
LOCUS	BE434257	591 bp	mRNA	EST	24-JUL-2000
DEFINITION	EST405335	tomato breaker fruit, TIGR Lycopersicon esculentum cDNA			
ACCESSION	BE434257	clone cLEG15H6, mRNA sequence.			
VERSION	BE434257.1	GI:9432100			
KEYWORDS	EST.				
SOURCE	tomato.				
ORGANISM	Lycopersicon esculentum				
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum; Lycopersicon.					
REFERENCE	1 (bases 1 to 591)				
AUTHORS	Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley				

S.D.		Generation of ESTs from tomato fruit tissue, breaker stage		
Unpublished (2000)				
Contact: David Frisch				
Clemson University Genomics Institute				
Clemson University				
100 Jordan Hall, Clemson, SC 29634, USA				
Tel: 864 656 4366				
Fax: 864 656 4293				
Email: dfrisch@clemson.edu				
5 prime sequence.				
Location/Qualifiers				
1..591				
/organism="Lycopersicon esculentum"				
/cultivar="TA496"				
/db_xref="taxon:4081"				
/clone="cLEG15H6"				
/clone.lib="tomato breaker fruit, TIGR"				
/tissue_type="pericarp"				
/dev_stage="breaker"				
/lab_host="SOLR"				
/note="Vector: pBluescriptSKmCvUadapt; Site_1: EcoRI; site_2: XhoI; fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."				
BASE COUNT	159 a	129 c	121 g	182 t
ORIGIN				
Query Match	7.2%;	Score 122.8;	DB 167;	Length 591;
Best Local Similarity	53.7%;	Pred. No. 6.5e-22;		
Matches 303;	Conservative 0;	Mismatches 252;	Indels 9;	Gaps 2;
QY	132	ATCCAGTCCCTCTCTGTTTACGACTTTTCGTCACCAAGAACACATTTCTTGGACACTGTT	191	
Db	7	ATAGGCGGATATATTTTACAGCTCTCCATTTTCAACCCGATTTTCGTTCAAAACATT	66	
QY	192	ATCCCTAAATCTTAAGCCCTCTTTGTCTCTCACTCTAAAACACTAGTCTCCGCTTAGCGGA	251	
Db	67	ATTCCCTCCTCTTAAAAATTCACCTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCCGA	126	
QY	252	AATTTGTTGATGCCGATCAAAATCGGGCGAAATGCGGAAGTTTCAGTACTCCGCTGTAGAG	311	
Db	127	AAGTTTGCTGTGCCACTAGATACAAACCGGATATCCTGAGTTACGTTA-----TGTGACA	180	
QY	312	GGCGACTCGATATACTTTGATTCGTTGGGAGTCTGACCAGGATTTTGTACTACTTAAAGGT	371	
Db	181	GGAGTTCTGTGCTGTACTTTTTTCGAGACTGATATGAATTTTCANATATCTCATTTGTT	240	
QY	372	CATCAACTGGTAGATTCCCAATGATTTGCATGGCCCTTTTTTATGTTATGCCAGGGTTATA	431	
Db	241	GACCATCCGGTAAGGCTAAGGATTTTAT---CACTTGTTCCTTAAGTTAGGGGAACCT	297	
QY	432	AGGACCATGCAAGACTATAAAGTGATCCCGCTCGTAGCCGTCGAAGTAACCGTTTTTCCCT	491	
Db	298	AAGGATGCACCGGGGTCCAACTAGCCCGCTTTAGCCACTTCAGGTGACACTTTTTTCCG	357	
QY	492	AACCGTGGCATACCCCTGGCTCTGACGGCACATCATTAATTCAGAGATGCTTAAAGATTTT	551	
Db	358	AATCTTGGTGATCCCATTTGGTTTCACTAAACATCATGTTGTTGGTATGGAGACTACTATA	417	
QY	552	GTAATGTTTCATCAATGCTTGGGCTATATTAACAAATTTGGGAAAGACGCGGACTTGTG	611	
Db	418	GCAGGGTTCAATTAAGCGGTGGGCTCTACTCCACAAATTCGGTGGACATGAACAATCTTA	477	
QY	612	TCCGCGAATCTTCTCCATCTTTTCGATAGATCGATAATCAAAAGATCTGTATGGCTAGAG	671	
Db	478	TCGNATGAGCTAATTCATTTTATGATAGTCCGTAGTAGTAAAGACCCATATGGCAAGGG	537	
QY	672	GAACATTTTGGAACGAATGCAA	695	
Db	538	ATGTCATCTCTGGAAGAAATGAAA	561	

S.D.	
Generation of ESTs from tomato fruit tissue, breaker stage	
Unpublished (2000)	
Contact: David Frisch	
Clemson University Genomics Institute	
Clemson University	
100 Jordan Hall, Clemson, SC 29634, USA	
Tel: 864 656 4366	
Fax: 864 656 4293	
Email: dfrisch@clemson.edu	
5 prime sequence.	
FEATURES	
Location/Qualifiers	
1..591	
/organism="Lycopersicon esculentum"	
/cultivar="TA96"	
/db_xref="taxon:4081"	
/clone="CLEG15H6"	
/clone_lib="tomato breaker fruit, TIGR"	
/tissue_type="Pericarp"	
/dev_stage="breaker"	
/lab_host="SOLR"	
/note="Vector: pBluescriptSKmCUadapt; Site_1: EcoRI;	
Site_2: XhoI; Fruit were harvested at the breaker stage	
(first sign of lycopene accumulation on the blossom end of	
the fruit). Fruit were cut in half and the seeds and	
locules were discarded prior to freezing the pericarp."	
BASE COUNT	159 a 129 c 121 g 182 t
ORIGIN	
Query Match 7.2%; Score 122.8; DB 167; Length 591;	
Best Local Similarity 53.7%; Pred. No. 6.5e-22;	
Matches 303; Conservative 0; Mismatches 252; Indels 9; Gaps 2;	
QY	132 ATGCAGTCCCTCTGTTTACGACTTTCGTAACCAAGACACATTTCTTGACACTGTT 191
Db	7 ATGACGGCGATATTATTATTTCAAGCTCTCCATTTTCAAAACCCGATTTCTGTTCAAAACATT 66
QY	192 ATCCCTAACTTAAGGCTCTTTGTCTCTCACTCTAAACACTACGTTCCGCTTAGCGGA 251
Db	67 ATTCCTCTCTTAAAAATTCACCTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCCGGA 126
QY	252 AATTTGTTGATGCCGATCAATCGGGCGAAATGCCGAAGTTTCAGTACTCCGCTTGATGAG 311
Db	127 AAGTTGCTTGTCCACTAGTACAAAGGGATATCCTGAGTTAGCTTA-----TGTGACA 180
QY	312 GGCGACTCGATAACTTTTGATTCGTTGGGAGTCTGACCCAGGATTTTGACTACCTTAAAGGT 371
Db	181 GGAGATTCCTGCTCTGTTACTTTTTCGAGACTGATGAATTTCAATTTCTCATTGTT 240
QY	372 CATCAACTGGTAGATTTCCAATGATTTGCGCTTTTATGTTATGCCACGGGTTATA 431
Db	241 GACCATCCGCTAAGGCTAAGGATTTTAT---CACTTTGTCTAAGTTAGGGGAACCT 297
QY	432 AGGACCATCAAGACTATAAAGTGTATCCGCTCTGAGCGTGCAGTAACCGTTTTCCT 491
Db	298 AAGGATGCACGGGGTCCCAACTAGCCCGCTCTTAGCCATTTCAGGTGACACTTTTCCG 357
QY	492 AACCGTGGCATAGCCGTGGCTCTGACGGACATCATTTCAATTCAGATGCTTAAAGTTT 551
Db	358 AATCTTGGTGATCCATTTGTTTTCACCAACCATCATGTTGTTGTTGATGAGCTACTATA 417
QY	552 GTAAATGTTTCATCAATGCTTGGGCTATATTAAACAATTTTGGGAAAGACCGGACTTGTG 611
Db	418 GCAGGTTTCATTAAAGCGTGGGCTCTACTCCCAAAATTCGGTGGACATGAACAATTTCTTA 477
QY	612 TCCGCAATCTTCTTCCATCTTTCGATAGATCGATAATCAAGATCTGTATGCCCTAGAG 671
Db	478 TCGAATGAGCTAATTCATTTTATGATAGTCCGTAGTAAAGACCCATATGGCAACAGG 537
QY	672 GAAACATTTTGGAAACGAATGCAA 695
Db	538 ATGTCCTCTGGGAAGAAATGAAA 561


```

Db 143 AGTACCAATTTCCAAACTCGATTTTCGTTCAAAACCATTTATTCCTACTCTTAACATTCAC 202
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 214 TGCTCTCTACTCTAAACACTAGCTTCCGCTTAGCGGAAATTTGTTGATGCCGATCAAT 273
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 203 TCTCCCTCACTCTCAAAACACTACACACCTTAGCTGGCAACATTTGCTTGCCACTAAAT 262
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 274 CGGGGAAATTCGCGAAGTTTCAGTACTCCGCTGATGAGGCGGACTCGATAAATTTGATCG 333
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 263 CGAGTGGTTATCTGAGTTGCGTTA-----TGTGACTGGAGATTCATATCTGTTACTT 316
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 334 TTCCGGAGTCTACACAGGATTTTGACTACTCTTAAAGGTCATCAACTGGTAGATTCOAATG 393
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 317 TTATTGAGACTGATATGAATTTCAATCATCTCATTTGGTAACCATCTCTCGAAATGCTAAGG 376
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 394 ATTTGATGCGCTTTTATGTTATCCACGGGTTATAGGACCATGCAAGACTATTAAG 453
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 377 ATTTTATCCCATCTCTCAATTTGGCAAC-----CTAAGGATGACCGGGGTCAAC 430
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 454 TGATCCCGCTCGTAGCGTGCAAGTAACCGTTTTCCTAACCGTGGCATAGCGTGGCTC 513
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 431 TAGTCCCGCTTAGCCATTCAGTGACACTTTTTCGGAATCTTGGCATATCCGTTGGTT 490
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 514 TGACGGACATCATTCATTCAGATGCTTAAAGTTTGTGAATGTTTCATCATCAATGCTTGGG 573
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 491 TTAGTAACCATCATCTCGCTTGTGTGGTGAATAACCATCGTGAATTCATAAGAACATGGG 550
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 574 CCTATTATACAAA 587
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 551 GTTTACTCAACAAA 564
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 7
BC097054 508 bp mRNA EST 29-JAN-2001
LOCUS EST461573 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION cSTB45P9 5' sequence, mRNA sequence.
ACCESSION BC097054
VERSION BC097054.1 GI:12587089
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE 1 (bases 1 to 508)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
source 1..508
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB45P9"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI. Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 127 a 130 c 87 g 164 t
ORIGIN
```

```

Query Match 6.5%; Score 111.4; DB 173; Length 508;
Best Local Similarity 54.0%; Pred. NO. 6.8e-19; Mismatches 226; Indels 9; Gaps 2;
Matches 276; Conservative 0;

Qy 36 GAAAAATGCCAAGTTACACACCATCTGACACAAAGATGTCGAGTATATGCGTACCGGTA 95
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 2 GAGCAATGCCAAGTTGCGCCACCTCCACGGGCGCAACAGAGCTACGCTCCCTCTA 61
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 96 ACATTCTCGATATCCCTGGTTGCACTTGAATAAGATGCAAGTCCCTCTGTTTATAGGAC 155
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 62 ACTATTATTTGATCATATGTTTGGTTCGATATGCGTCGGATTTTATTTACTACAAG 121
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 156 TTTCCGTAACCCAAAGACACATTTCTTGGACACTGTTATCCCTAATCTTAAGGCTCTCTTG 215
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 122 CTACCAATTTCCAAACTCGATTTTCGTTCAACCATTTATTCCTACTCTTAACATTCAC 181
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 216 TCTCTCACTCTAAACACTACGTTTCGCTTAGCGGAAATTTGTTGATGCCGATCAATCG 275
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 TCCTCACTCTCAACACTACACACCTTAGCTGGCAACATTCCTTGTCCACTAAATCG 241
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 276 GCGAAATGCCAAGTTTCAGTACTCCGCTGATGAGGGGACATCGGATGATGCTGTT 335
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 242 AGTGGTTATCTCGATTCGCTTA-----TGTGACTGGAGATTCATATCTGTACTTTT 295
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 336 GCGGAGTCTGACACGAGTTTGGACTACCTTAAAGGTCATCAACTGGTAGATTTCCAATGAT 395
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 296 ATTGAGACTGATATGAATTTCAATCATCTCATTTGGTAACCATCTCTCGAAATGCTAAGGAT 355
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 396 TTGCATGGCCCTTTTATGTCACCGGGTTATAGGACCATGCAAGACTATAAAGTG 455
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 356 TTTTATCCCTTCATTCCTCAATTTGGCACAACCTAAGGATCGACCGGGGTC---AAACTA 412
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 456 ATCCCGCTCGTAGCCGTGAAGTAACCGTTTTTCTTAACCGTGGCATAGCGGTGCTG 515
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 413 GTCCCGCTCTTAGCCATTCAAAGTGACACTTTTTCGGAATCTTGGCATATCCGTTGGTTT 472
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 516 ACGGCACATCATTCATTCAGATGCTAATAA 546
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 473 AGTAACCATCATCTCGCTTGTGATGAATA 503
      | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 8
BC051101 578 bp mRNA EST 16-OCT-2000
LOCUS EST436276 tomato developing/immature green fruit Lycopersicon
DEFINITION esculentum cDNA clone cLEM21H7 5' sequence, mRNA sequence.
ACCESSION BC051101
VERSION BC051101.1 GI:10804997
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE 1 (bases 1 to 578)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,R.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, immature green
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.
FEATURES
source 1..578
Location/Qualifiers
```



```
Query Match          6.3%; Score 107.6; DB 118; Length 603;
Best Local Similarity 57.2%; Pred. No. 7.3e-18;
Matches 263; Conservative 0; Mismatches 179; Indels 18; Gaps 3;

Qy 963 GATTGCCGAGACTTCTGACGCCGCCGCTGCTACTACTTTGGCAACTGCTTGGC 1022
Dy 602 GATTTCAGGCGGATTCATCAATCCACCTCTCCTCAATCTTATTTTGGAAATGTCATGTT 543
Qy 1023 TCATCCGTTGCAAAAGCAACA--CATAAAGACTAGTTGGGATAAAGGCTCTTGT 1079
Dy 542 GGGTATGTCACAAAATCAACAAGGATGTTGACTAGTTGGAAAGGAGGCTTTAAAAAT 483
Qy 1080 GCAGTTGCAGCTATTGGAGAGACCATCAAAAAGAGTTGCCACAACGAAAGGCTTCTT 1139
Dy 482 CGGTAGAAATCAATGGAGAGATCATTCACGAAAAAATGAAGGATGATGATGGTCTT 423
Qy 1140 GCAGATGCAAAACCTGGTTATTCGGAATCTTAATGGAATCCCTTCAAAAAGATTTCTCGG 1199
Dy 422 AATGGTG-----ATTGGTTAAAAGTATTAGACAACATAGATGTGATTCGATCATTTTCA 369
Qy 1200 ATTACCGGATCGCTAAGTTTCGATTCGATGCTGATGATTTTGGATGGGAAAGCCTGCA 1259
Dy 368 ATTGCTGATCGCCAAAACATGACTTATATCTGCTGCTGATTTTGGATGGGAGAGCCGCA 309
Qy 1260 AAATTTGCATTTACCTC-----TGTTGATTATGAGAAATTCATTTATGTTGATTCAG 1310
Dy 308 AAGTTAGAATTTATTTTCATCGACAATGATGTTGTTGGAATTTCAATGCTCTTAGTAAA 249
Qy 1311 TCCAGGGATTTTGAAMAGGTGTGGAGATTCGATGATTCATTCGCTCAAGATTCATATGAT 1370
Dy 248 TCTAAGGATTTTGTATGGAGATTTAGAGATTTGCTGTTGTTGTTCTCTAAACCTCGAATGA 189
Qy 1371 GCATTTGCAAAATCTTTGAAGAAGGCTTTTTCCTTTGT 1410
Dy 188 GCTTTTGTACTATATTCACCTCCCGGCTTAGCTTTCTAT 149

RESULT 12
LOCUS BE436185 542 bp mRNA EST 24-JUL-2000
DEFINITION EST407263 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG31A19, mRNA sequence.
ACCESSION BE436185
VERSION BE436185.1 GI:9434028
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 542)
AUTHORS Alcalá, J., Vrebalov, J., White, R., van der Hoeven, R.S., Holt, I.E.,
Liang, F., Hansen, T.S., Craven, M.B., Bowman, C.D., Ronning, C.M.,
Nierman, W., Fraser, C.M., Martin, G.B., Giovannoni, J.J. and Tanksley
, S.D.
Generation of ESTs from tomato fruit tissue, breaker stage
Unpublished (2000)
CONTACT: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
LOCATION/Qualifiers
1. .542
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG31A19"

FEATURES
source
1. .542
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCUadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopenic accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 175 a 70 c 132 g 165 t
ORIGIN
```

Query Match 6.3%; Score 107.2; DB 167; Length 542;
Best Local Similarity 56.7%; Pred. No. 9e-18;
Matches 262; Conservative 0; Mismatches 188; Indels 12; Gaps 3;

Qy 833 AATACGTTAAACGAGTTCAACATGAGTGTGGATACGATGACATGCGTCAAAATC 892
Dy 68 AACTCATGTAACATCTTTTACTGTAACGAGTGCTATATATGGAATGTTGATAAAATC 127
Qy 893 AAAAGATCAGCTCTATCAGAGGATCATCAACGACGAAATGAGCTCGAGTACTTCAG 952
Dy 128 ---AGAGACCGGATAGAGAGAGATGATCATAGATGATAGTGTAAATGGAATGTTTCGG 184
Qy 953 TTTTACAGCGGATGCCGAGGACTTCTGACGCCGCCGCTGTCGCCCTAACTACTTTGGCAA 1012
Dy 185 ATGTGACGACATTTAGAGCGCGATTCATCAATCCACCACCTCTCTCAATCTTATTTGGGAA 244
Qy 1013 CTGCTTCGCGTATCGGTTGCCAAAGCAACA---CATAAAGAGTTAGTTGGGGATAAAGG 1069
Dy 245 TTGCATAGTTGGGTATGTCACAAAATCAATAAGGACTTTGACTTAGTTGGAAGGAAAGG 304
Qy 1070 GCTTCTTGTTCAGTTGCGAGCTATTGGAGAGCCATTCAAAAGAGTTGTCACACGAAAA 1129
Dy 305 GTTTAAATTCGGGTAGAAATCAATTGGAGAAAGTCAATCAAGAAAAATGAAGGATGATGA 364
Qy 1130 AGCGTCTTTCGAGATGCAAAAACCTTGTTATTCGGAATCTTAATGGAATCCCTTCAAAAAG 1189
Dy 365 ATGGTCTCTTAATGGTG-----ATTGGTTAAAGATATACGACAATGTAGATGTGATTCG 418
Qy 1190 ATTTCTCGGGATTAACCGGATCGCTTAAGTTCGATTCGATGCTGATGTTGATGTTGGATGGG 1249
Dy 419 ATCATTTTCAATTCGTGATCGCAAAAACATGACTTATATGCTGCTGATTTTGGATGGG 478
Qy 1250 AAGCCTGCAAAATTTGACATTACCTCTGTTGATATATCCAGA 1291
Dy 479 AAGAGCCGCAAGTTGGAATTCATTTCCATTGACAATGATGA 520

RESULT 13
LOCUS BG598154 659 bp mRNA EST 12-APR-2001
DEFINITION EST496832 cSTS Solanum tuberosum cDNA clone cSTS20M23 5' sequence,
mRNA sequence.
ACCESSION BG598154
VERSION BG598154.1 GI:13616294
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 659)
AUTHORS van der Hoeven, R., Bezzerides, J., Sun, H., Cho, J., Chiemlingo, A.,
Bougril, O., Buell, C.R., Ronning, C., Tanksley, S. and Baker, B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
CONTACT: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

SOURCE ORGANISM		Lycopersicon esculentum	
REFERENCE AUTHORS		Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F., Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M., Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.	
TITLE		Generation of ESTs from tomato fruit tissue	
JOURNAL COMMENT		Unpublished (1999)	
CONTACT		David Frisch	
Clemson University Genomics Institute		Clemson University	
100 Jordan Hall, Clemson, SC 29634, USA		Tel: 864 656 4366	
Fax: 864 656 4293		Email: dfrisch@CLEMSON.EDU	
5 prime sequence.		Location/Qualifiers	
FEATURES		1. .497	
source		/organism="Lycopersicon esculentum"	
		/cultivar="TA496"	
		/db_xref="taxon:4081"	
		/clone="cLEF42M23"	
		/clone_lib="tomato fruit mature green, TAMU"	
		/tissue_type="fruit pericarp"	
		/dev_stage="mature green (3-5 days pre-ripening)"	
		/lab_host="SOLR"	
		/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"	
BASE COUNT	132 a	106 c	109 g 150 t
ORIGIN			
Query Match 6.0%; Score 101.6; DB 122; Length 497;			
Best Local Similarity 54.0%; Pred. No. 2.7e-16;			
Matches 256; Conservative 0; Mismatches 209; Indels 9; Gaps 2;			
QY	210	TCCTTTGTCCTCACTCTAAACACACTAGCTTCGCTTAGCGGAAATTTGTTGATCGCGATC	269
Db	2	TCACCTCTCCCTCACTCTCAACACTATACGCCCTTAGCCGGAACGTTGCTGTGCCACTA	61
QY	270	AAATCGGGCGAAATGCCGAAGTTTCAGTACTCCCGTATGAGGCGGACTCGATAACTTTG	329
Db	62	GATACAAACGGGATATCCTGAGTTACGTTA-----TGTGACAGGAGATTCTGTCTGT	115
QY	330	ATCGTTGCGGAGTCTGACCAAGGATTTTGACTACCTTAAAGGTCACTCAACTGGTAGATTC	389
Db	116	ACTTTTTTCGAGACTGATATGAATTTCAATTTATCTCATTTGTTGACCATCCGCGTAAGGCT	175
QY	390	AATGATTTGCAATGGCCTTTTATGTTATGCCACGGGTTATAAGGACCATGCAAGACTAT	449
Db	176	AAGATTTTTTATCACTTTGTTC---TAAGTTAGGGGAACCTAAGGATGACCGGGGGTC	232
QY	450	AAAGTGATCCCGCTCGTAGCGTGAAGTAACCGGTTTTTTCCTAACCGTGGCATAGCCGTG	509
Db	233	CAACTAGCCCGCTCTTAGCCATTCAGGTGACACTTTTTCGGAATCTTGGTGTATCCATT	292
QY	510	GCTCTGACGCGACATCATCAATTTGCGAGATGCTTAAAGTTTGTAAATGTTTCATCAATGCT	569
Db	293	GGTTTCACTAACCATCATGTGTGTTGATGAGGACTACTATAGCAGGGTTCATTAAAGCG	352
QY	570	TGGGCTATATTAAACAATTTGGGAAGACGCGGACTTGTGTCGCGAATCTTCTTCCA	629
Db	353	TGGGCTCTACTCCACAATTTCCGGTGGACATGAACAATTTCTTATCGAATGAGCTAATCCA	412
QY	630	TCTTTCGATAGATCGATAATCAAAAGATCTGTATGGCCTAGAGGAACATTTTGG	683

FEATURES		Location/Qualifiers	
source		1. .659	
		/organism="Solanum tuberosum"	
		/cultivar="Kennebec"	
		/db_xref="taxon:4113"	
		/clone="cSTS20M23"	
		/clone_lib="cSTS"	
		/tissue_type="sprouting eyes from tubers"	
		/dev_stage="12-14 weeks post harvest"	
		/lab_host="SOLR"	
		/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Various sizes of sprouting eyes (2mm to 15mm) were taken from tubers. The tubers were incubated at 26C in the dark for 2-3 weeks prior to sprouting. The eyes were frozen in liquid nitrogen immediately upon removal from tubers."	
BASE COUNT	178 a	163 c	112 g 206 t
ORIGIN			
Query Match 6.0%; Score 102.8; DB 155; Length 659;			
Best Local Similarity 53.4%; Pred. No. 1.4e-16;			
Matches 310; Conservative 0; Mismatches 257; Indels 13; Gaps 4;			
QY	34	TTGAAAAATGCCAAGTTACACCACTCTCACAACAGATGTCGAGTTATCGCTACCGG	93
Db	90	TTGAGCAATGCCAAGTTGCCCACTCCCAACGGCGGCACAGAGCTAACGCTCCCTC	149
QY	94	TACATCTTCGATATCCCGTGGTTCGATCTGAATAGATGACAGTCCCTTCGTTTACG	153
Db	150	TAACTTATTTGATCT-ATATGGTTCGGTTCGCTATATGCTCGGATTTTATCTACA	208
QY	154	ACTTTCGTTACCAACACATTTCTTGACACTCTTATCCCTATCTTAAGCCCTCTT	213
Db	209	AGTACCAATTTCCAACTCGATTCGTTCAAAACCAATTTCTCTACTCTTAAACATTCAC	268
QY	214	TGTCTCTCACTCTAAACACTAGCTTCCGCTTACGCGAAATTTGTTGATGCCGATCAAAAT	273
Db	269	TCCTCCCTCACTCTCAACACTACACACCCTTAGCTGGCAACATTTGCTGTCCACTAAAT	328
QY	274	CGGGCAATCGGAATTCAGTACTCCGCTGATGAGGCGGACTCGATAACTTTGATCG	333
Db	329	CGAGTGGTTATCTGATGTCGGTTA-----TGTGACTGGAGATCTATATCTGTACTT	382
QY	334	TTCCGGAGTCTGACCAAGATTTTGACTACCTTAAAGTCACTCACTGGTAGATTCCAATG	393
Db	383	TTATGAGACTGATAGAAATTTCAATCATCTCATTTGTAACCATCTCTCGAAATGCTAAG	442
QY	394	ATTGCAATGGCCCTTTTATGTTATGCCACGGGTTATAAGGACCATGCAAGACTATAAG	453
Db	443	ATTTTATCCCTTCATCTCTCAATTTGGCACAACTAAGGATGACCGGGGTC--AAAC	499
QY	454	TGATCCCGCTGATGCGGTCGAAGTAACCGTTTTCCTAACCGTGGCAGTACCGGT--GG	510
Db	500	TAGTCCCGCTCTTAGCCATTTCAAGTGACACTTTTTCGGAATCTTGCCATATCCCGTTGT	559
QY	511	CTCTGACGCGACATCATCAATTTGACAGATGCTAAAGTTTTGAATGTTTCATCAATGCTT	570
Db	560	GTTTACTAACCCATCATGCGCTTGTGATGGAATACCATCGTGAATTCATAGAACAAT	619
QY	571	GGGCTATATTAAACAATTTGGGAAAGACCGGACTTGT 610	
Db	620	GGGTTTACTCAACAATTTCCGCGGTGATGAACAGTGCTT 659	
RESULT 14			
AW930934			
LOCUS AW930934 497 bp mRNA EST 30-MAY-2000			
DEFINITION EST356777 tomato fruit mature green, TAMU Lycopersicon esculentum			
ACCESSION AW930934			
VERSION AW930934.1 GI:8106335			
KEYWORDS EST.			

THIS PAGE BLANK (USPTO)

Db 2260 AATGTATACCTTTCACATGCTCTTTGTGCTTTTACAAATAAATGATATGAAC 2319
QY 1676 TGTAAAAA 1703
Db 2320 TCAAAAAA 2347

RESULT 5
US-08-726-160-9
; Sequence 9, Application US/08726160
; Patent No. 5734016
; GENERAL INFORMATION:
; APPLICANT: LEVENS, DAVID L., DUNCAN,
; APPLICANT: ROBERT C., AND AVIGAN, MARK I.
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 24
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/726.160
; FILING DATE: 04-OCT-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/021.608
; FILING DATE: 22-FEB-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26.728
; REFERENCE/DOCKET NUMBER: 2026-4063US1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2381
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: No
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; ORGANISM: Human
; CELL LINE: HL60
; FEATURE:
; OTHER INFORMATION: 470 bp variable
; OTHER INFORMATION: region where R is A or G.
US-08-726-160-9

Query Match 2.3%; Score 38.4; DB 1; Length 2381;
Best Local Similarity 64.8%; Pred. No. 0.22;
Matches 57; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

QY 1616 ATGTATAATGCCATTATATCTTCCATAAAGTATCTTCAATAGAGAACAATGTTATG 1675
Db 2260 AATGTATACCTTTCACATGCTCTTTGTGCTTTTACAAATAAATGATATGAAC 2319
QY 1676 TGTAAAAA 1703

Db 2320 TCAAAAAA 2347
RESULT 6
PCT-US94-01782-9
; Sequence 9, Application PC/TUS9401782
; GENERAL INFORMATION:
; APPLICANT: THE GOVERNMENT OF THE UNITED STATES
; APPLICANT: AS REPRESENTED BY THE SECRETARY, DEPARTMENT OF
; APPLICANT: HEALTH AND HUMAN SERVICES
; TITLE OF INVENTION: NOVEL FUSE BINDING
; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MORGAN & FINNEGAN
; STREET: 345 PARK AVENUE
; CITY: NEW YORK
; STATE: NEW YORK
; COUNTRY: USA
; ZIP: 10154
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY DISK
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: WORDPERFECT 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/01782
; FILING DATE: 22-FEB-1994
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: U.S. 08/021.608
; FILING DATE: 22-FEB-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: WILLIAM S. FEILER
; REGISTRATION NUMBER: 26.728
; REFERENCE/DOCKET NUMBER: 2026-4063PCT
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 758-4800
; TELEFAX: (212) 751-6849
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2381
; TYPE: Nucleic acid
; STRANDEDNESS: Double
; TOPOLOGY: Unknown
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: No
; ANTI-SENSE: YES
; ORIGINAL SOURCE:
; ORGANISM: Human
; STRAIN:
; INDIVIDUAL ISOLATE:
; DEVELOPMENTAL STAGE:
; HAPLOTYPE:
; TISSUE TYPE:
; CELL TYPE:
; CELL LINE: HL60
; ORGANELLE:
; POSITION IN GENOME:
; CHROMOSOME/SEGMENT:
; MAP POSITION:
; UNITS:
; FEATURE:
; NAME/KEY:
; LOCATION:
; IDENTIFICATION METHOD:
; OTHER INFORMATION: 470 bp variable
; OTHER INFORMATION: region where R is A or G.
PCT-US94-01782-9

Query Match 2.3%; Score 38.4; DB 5; Length 2381;
Best Local Similarity 64.8%; Pred. No. 0.22;

Matches 57; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 1616 ATTGTATAATGCCATTATATCTTCCATAAAGTATCTCTATGCAATAGAACATGTTATG 1675

Db 2260 AATGTATATCTTCACTTCCAAATGCCCTGTTTGTGCTTTACAATAAATGATATGAACC 2319

Qy 1676 TGTAAAAAIAAAAAAAAAAAAAAAAAAAAA 1703

Db 2320 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2347

RESULT 7

US-08-021-608D-1

; Sequence 1, Application US/08021608D

; Patent No. 5580760

; GENERAL INFORMATION:

; APPLICANT: LEVENS, DAVID L., DUNCAN,

; APPLICANT: ROBERT C., AND AVIGAN, MARK I.

; TITLE OF INVENTION: NOVEL FUSE BINDING

; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/021,608D

; FILING DATE: 22-FEB-1993

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER:

; FILING DATE:

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: WILLIAM S. FEILER

; REGISTRATION NUMBER: 26,728

; REFERENCE/DOCKET NUMBER: 2026-4063

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2384

; TYPE: Nucleic acid

; STRANDEDNESS: Double

; TOPOLOGY: Unknown

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: No

; ANTI-SENSE: YES

; ORIGINAL SOURCE:

; ORGANISM: Human

; CELL LINE: HL60

; FEATURE:

; OTHER INFORMATION: 473 bp

; OTHER INFORMATION: variable region where R is A or G.

US-08-021-608D-1

Query Match 2.3%; Score 38.4; DB 1; Length 2384;

Best Local Similarity 64.8%; Pred. No. 0.22;

Matches 57; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 1616 ATTGTATAATGCCATTATATCTTCCATAAAGTATCTCTATGCAATAGAACATGTTATG 1675

Db 2263 AATGTATATCTTCACTTCCAAATGCCCTGTTTGTGCTTTACAATAAATGATATGAACC 2322

Qy 1676 TGTAAAAAIAAAAAAAAAAAAAAAAAAAAA 1703

Db 2323 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2350

RESULT 8

US-08-726-160-1

; Sequence 1, Application US/08726160

; Patent No. 5734016

; GENERAL INFORMATION:

; APPLICANT: LEVENS, DAVID L., DUNCAN,

; APPLICANT: ROBERT C., AND AVIGAN, MARK I.

; TITLE OF INVENTION: NOVEL FUSE BINDING

; TITLE OF INVENTION: PROTEIN AND CDNA THEREFOR

; NUMBER OF SEQUENCES: 24

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: MORGAN & FINNEGAN

; STREET: 345 PARK AVENUE

; CITY: NEW YORK

; STATE: NEW YORK

; COUNTRY: USA

; ZIP: 10154

; COMPUTER READABLE FORM:

; MEDIUM TYPE: FLOPPY DISK

; COMPUTER: IBM PC COMPATIBLE

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: WORDPERFECT 5.1

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/726,160

; FILING DATE: 04-OCT-1996

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 08/021,608

; FILING DATE: 22-FEB-1993

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: WILLIAM S. FEILER

; REGISTRATION NUMBER: 26,728

; REFERENCE/DOCKET NUMBER: 2026-4063US1

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (212) 758-4800

; TELEFAX: (212) 751-6849

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 2384

; TYPE: Nucleic acid

; STRANDEDNESS: Double

; TOPOLOGY: Unknown

; MOLECULE TYPE: cDNA

; HYPOTHETICAL: No

; ANTI-SENSE: YES

; ORIGINAL SOURCE:

; ORGANISM: Human

; CELL LINE: HL60

; FEATURE:

; OTHER INFORMATION: 473 bp

; OTHER INFORMATION: variable region where R is A or G.

US-08-726-160-1

Query Match 2.3%; Score 38.4; DB 1; Length 2384;

Best Local Similarity 64.8%; Pred. No. 0.22;

Matches 57; Conservative 0; Mismatches 31; Indels 0; Gaps 0;

Qy 1616 ATTGTATAATGCCATTATATCTTCCATAAAGTATCTCTATGCAATAGAACATGTTATG 1675

Db 2263 AATGTATATCTTCACTTCCAAATGCCCTGTTTGTGCTTTACAATAAATGATATGAACC 2322

Qy 1676 TGTAAAAAIAAAAAAAAAAAAAAAAAAAAA 1703

Db 2323 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2350


```

:
:
: CLASSIFICATION:
:
: ATTORNEY/AGENT INFORMATION:
:
:   NAME:  Verser, Carol Talkington
:
:   REGISTRATION NUMBER:  37,459
:
:   REFERENCE/DOCKET NUMBER:  HW-5-C1
:
: TELECOMMUNICATION INFORMATION:
:
:   TELEPHONE:  970/493-7272
:
:   TELEFAX:  970/484-9505
:
: INFORMATION FOR SEQ ID NO.: 102:
:
:   SEQUENCE CHARACTERISTICS:
:
:     LENGTH:  184 nucleotides
:
:     TYPE:  nucleic acid
:
:
:

```

US-09-065-474-103

QY 1042 CACATAAAGAGTTAGTTGGGGATAAAGGCTTCTTGTTCAGTTGCACGTATTGGAGAAG 1101

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 13:54:25 ; Search time 445.49 seconds
(without alignments)
2400.315 Million cell updates/sec

Title: US-08-894-356C-1

Perfect score: 1703

Sequence: 1 TCATTGAGCAATCCAA.....AAAAAAAAAAAAAAAAAAAAA 1703

Scoring table: IDENTITY_NUC

Gapop 10.0, Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N_Geneseq_0601.*

- 1: /SIDSL/gcgdata/geneseq/geneseq/NA1980.DAT.*
- 2: /SIDSL/gcgdata/geneseq/geneseq/NA1981.DAT.*
- 3: /SIDSL/gcgdata/geneseq/geneseq/NA1982.DAT.*
- 4: /SIDSL/gcgdata/geneseq/geneseq/NA1983.DAT.*
- 5: /SIDSL/gcgdata/geneseq/geneseq/NA1984.DAT.*
- 6: /SIDSL/gcgdata/geneseq/geneseq/NA1985.DAT.*
- 7: /SIDSL/gcgdata/geneseq/geneseq/NA1986.DAT.*
- 8: /SIDSL/gcgdata/geneseq/geneseq/NA1987.DAT.*
- 9: /SIDSL/gcgdata/geneseq/geneseq/NA1988.DAT.*
- 10: /SIDSL/gcgdata/geneseq/geneseq/NA1989.DAT.*
- 11: /SIDSL/gcgdata/geneseq/geneseq/NA1990.DAT.*
- 12: /SIDSL/gcgdata/geneseq/geneseq/NA1991.DAT.*
- 13: /SIDSL/gcgdata/geneseq/geneseq/NA1992.DAT.*
- 14: /SIDSL/gcgdata/geneseq/geneseq/NA1993.DAT.*
- 15: /SIDSL/gcgdata/geneseq/geneseq/NA1994.DAT.*
- 16: /SIDSL/gcgdata/geneseq/geneseq/NA1995.DAT.*
- 17: /SIDSL/gcgdata/geneseq/geneseq/NA1996.DAT.*
- 18: /SIDSL/gcgdata/geneseq/geneseq/NA1997.DAT.*
- 19: /SIDSL/gcgdata/geneseq/geneseq/NA1998.DAT.*
- 20: /SIDSL/gcgdata/geneseq/geneseq/NA1999.DAT.*
- 21: /SIDSL/gcgdata/geneseq/geneseq/NA2000.DAT.*
- 22: /SIDSL/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	ID	Description
1	1698.2	99.7	1703	17 AAT37308
2	202.4	11.9	1508	17 AAT37312
3	193	11.3	1518	17 AAT37313
4	186	10.9	1622	17 AAT37309
5	172.2	10.1	1479	17 AAT37311
6	110	6.5	936	22 AAF58252
7	110	6.5	936	22 AAF58254
8	110	6.5	936	22 AAF58257
9	110	6.5	936	22 AAF58259
10	110	6.5	936	22 AAF58262
11	110	6.5	938	22 AAF58255

12	106.8	6.3	936	22 AAF58252	Oligonucleotide D1
13	106.8	6.3	936	22 AAF58254	Oligonucleotide D1
14	106.8	6.3	936	22 AAF58257	Oligonucleotide D1
15	106.8	6.3	936	22 AAF58259	Oligonucleotide D2
16	106.8	6.3	936	22 AAF58262	Oligonucleotide D2
17	106.8	6.3	938	22 AAF58255	Oligonucleotide D1
18	81	4.8	1648	21 AAC64785	Lemon acyl transferase
19	53.6	3.1	1379	21 AAC64852	Arabidopsis thaliana
20	45.4	2.7	1502	21 AAC49421	Arabidopsis thaliana
21	45.4	2.7	1504	21 AAC33536	Arabidopsis thaliana
22	45	2.6	244	22 AAF58238	Oligonucleotide D1
23	44.2	2.6	244	22 AAF58238	Oligonucleotide D1
24	43	2.5	1240	21 AAC53227	Arabidopsis thaliana
25	42.4	2.5	1486	21 AAC98137	Human colon cancer
26	41.8	2.5	2222	21 AAA27236	Human MEK1 DNA
27	40	2.3	408	22 AAF17559	Human breast cancer
28	40	2.3	2691	21 AAC79685	Pinus radiata cell
29	39.8	2.3	959	22 AAC66940	Proteasome subunit
30	39.8	2.3	3581	15 AAQ44391	Sequence of murine
31	39	2.3	340	20 AAX51737	DNA encoding a hum
32	39	2.3	1373	19 AAV52969	Human G1 protein 1
33	39	2.3	1701	21 AAZ33336	Human secreted pro
34	39	2.3	5713	21 AAC76360	Human ORFX ORF1915
35	39	2.3	7122	22 AAF31100	Methionine synthas
36	39	2.3	7122	22 AAC91207	Human methionine s
37	38.8	2.3	1544	21 AAA46483	cDNA encoding a fi
38	38.8	2.3	1670	22 AAF24992	Nucleotide sequenc
39	38.6	2.3	932	21 AAZ51920	N. tabacum TOB-F12
40	38.6	2.3	1106	21 AAA29167	Rice 4-hydroxyphen
41	38.6	2.3	4292	21 AAF21792	Human breast and o
42	38.4	2.3	2381	15 AAO68910	Human c-myc far up
43	38.4	2.3	2384	15 AAO68909	Human c-myc far up
44	38.2	2.2	1570	21 AAC76262	Human ORFX ORF1817
45	38.2	2.2	1820	20 AAZ00434	Human secreted pro

ALIGNMENTS

RESULT 1

ADT37308
ID AAT37308 standard; cDNA to mRNA; 1703 BP.

AC AAT37308;

DT 06-FEB-1997 (first entry)

XX Aromatic acyl transferase coding sequence.

DE Aromatic acyl transferase: transformation; anthocyanin pigment;

KW plants; acylation; colour; tone; colouration; colour change;

KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;

KW Scenecio cruentus; Lavandula angustifolia; ds.

OS Gentiana triflora var. japonica (Clone pcAT4).

FH Key Location/Qualifiers

FT CDS 6..1415

FT /*tag= a

FT /product= Aromatic acyl transferase.

XX WO9625500-A1.

XX 22-AUG-1996.

XX 16-FEB-1996; 96WO-JP00348.

XX 30-JAN-1996; 96JP-0046534.

XX 17-FEB-1995; 95JP-0067159.

XX 29-JUN-1995; 95JP-0196915.

XX (SUNR) SUNTORY LTD.

PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
PI Nakao M, Tanaka Y, Yonekura K;
XX WPI: 1996-393401/39.
DR P-PSDB: AAW04722.
XX
PT DNA coding for aromatic acyl transferase - for transforming plants
PT which produce anthocyanin pigments and thus altering colour tone,
PT e.g. of flowers
XX
PS Claim 4; Page 53-57; 94pp; Japanese.
XX
CC Vectors containing DNA fragments encoding proteins of plant origin
CC with aromatic acyl transferase activity may be used to transform
CC plants which produce anthocyanin pigments. The aromatic acyl
CC transferase acylates the pigments in the flower resulting in colour
CC tone changes and allowing new colourations to be produced. Six
CC specific DNA sequences encoding aromatic acyl transferase from
CC different plants are described in AAT37308-T37313.
XX
SQ Sequence 1703 BP; 512 A; 353 C; 356 G; 482 T; 0 other;

Query Match 99.7%; Score 1698.2; DB 17; Length 1703;
Best Local Similarity 99.8%; Pred. No. 0;
Matches 1700; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 TCATTATGGAGCAATCCAAATGTTGAGGTTCTTGAAAATGCCAAGTTACACCACCAT 60
DB 1 tcattatggagcaaatccaaatggtgaggtcttgaaaaatgccaaagtacaccacac 60
QY 61 CTGACAAACAGATGTCAGTTATCGTACCGGTAAACATCTTCGATATCCCGTGTTC 120
DB 61 ctgacaaacagatgtcagttatcggtacccggttaaacatcttcgatatccctggtgc 120
QY 121 ACTTGAATAAGATCAGTCCCTCTGTTTACGACTTTCCGTACCCCAAGAACACATTCT 180
DB 121 acttgaataagatcagtcctctctgtttacgactttccgtaccccaagaacacattct 180
QY 181 TGGACACTGTTATCCCAATCTTAAGCCCTTTGTCTCTCAGTCTAAACACTACGTTTC 240
DB 181 tggacactgttatcccaatcttaagccctttgtctctcactctaaacactaccttc 240
QY 241 CGCTTAGCGGAAATTTGTTGATCGCGTCAATCGCGGAAATCGCGAAGTTTCAGTACT 300
DB 241 cgccttagcggaaatttgttgatcgcgatcaaatcgcggaatcgcggaagtctcagtact 300
QY 301 CCCGTGATAGGGCGACTCGATACCTTTGATCGGTGCGGAGTCTGACCAGGATTTGACT 360
DB 301 cccgtgataggcgactcgatacctttgatcggtgcggaatcgcggaagtctcagtact 360
QY 361 ACCTTAAGGTCATCACTGGTAGTCCCAATGATTGATGCGCTTTTATGTTATGC 420
DB 361 accttaaggtcatacctggtagtcccaatgattgatcgctttttatgtatgc 420
QY 421 CACGGGTTATGAAGCAGCATCAAGCTATAAGTGATCCCGCTCGTAGCCGTGCAAGTAA 480
DB 421 cacgggttataagcagcatcaagctataaagtga lcccgctcgtagccgtgcaagtaa 480
QY 481 CCGTTTTTCTTAACCGGATACCGCTGGGCTCTGAGGCGACATCAATTAATTCAGATG 540
DB 481 ccgtttttcttaacccggataccgctggctctgagcgacacatcatcattcaatgcatg 540
QY 541 CTAAGAGTTTGTAAATGTTCAATGCTTGGGCTATATTAACAAATTTGGGAAAGACG 600
DB 541 ctaagagtttgttaagtctcatcaatgcttgggctataatacaaatttgggaaagacg 600
QY 601 CGGACTTGTTCGCGAAATCTTCTCCATCTTTCGATAGATGATCAATCAAGATCTGT 660
DB 601 cggacttgttgcggaatcttctccatcttcgatatagatgataatcaaatgcatgt 660
QY 661 ATGCCCTAGAGGAACATTTTGGAAACGAAATGCAAGATGTTCTTGAATGTTCTCTAGAT 720
DB 661 atgcccttagaggaaacatTTTGGAAACGAAATGCAAGATGTTCTTGAATGTTCTCTAGAT

DB 661 atggccttagaggaacacattttgggaacgaatgcaagatgttcttgaaatgttctctagat 720
QY 721 TTGGAAGCAAAACCCCTCGATTCAACAGGTACGAGCTACATATGCTCTCCCTTCCTG 780
DB 721 ttggaagcaaaacccctcgcattcaacaggttacgaggtacatagtctctctccctgtg 780
QY 781 AAATCCAGAAAGCTAAAGAACAAAGTACTGAATCTCAGAGGATCCGAACCCGACAAATAGCTG 840
DB 781 aaatccagaagctaaagaacaaagtactgaatctcagaggtccgaaccgacacacgtg 840
QY 841 TAACGACGTTTCAATGACGTTGCGGATGACGTATGACATGATGATGATCAATCAAAAGATG 900
DB 841 taacgacgtttcaaatgacgtgtgtgacgtatgacatgcatggtcgaatcaaaagatg 900
QY 901 ACGTCGATCAGAGGAATCATCGAACGACGAAATAGCTCGAGTACTTCAAGTTTACAG 960
DB 901 acgtcgatcagaggaatcatcgaaacgacgaaatagctcgagcacttcagttttacag 960
QY 961 CGGATTGCCGAGGACTTCTGACGCCGCCGCTGTCGCCCTTAATCTTTGGCAACTGCTCTG 1020
DB 961 cggattgccgaggaacttctgacgccccggtgcgcctaaactactttggcaactgtctg 1020
QY 1021 CGTCATCGCTTGCAAAAGCAACACATTAAGAGTTAGTTGGGGATAAAGGCTTCTTGTG 1080
DB 1021 cgtcatcgcttgcaaaagcaacacataaagagttagttgggataaaagggtcttctgtg 1080
QY 1081 CAGTTGCGAGCTATTGGAGAACCCATTGAAAAGAGTTGCAACAGAAAGGCGCTTCTTG 1140
DB 1081 cagttgagctattggagaaagccattgaaagaggttgcaacacgaaagggcgttcttg 1140
QY 1141 CAGATGCAAAACATCTGGAATCTAATGGAATCCCTTCAAAAGATTTCTTCGGGA 1200
DB 1141 cagatgcaaaacttggttatcggaatcctaaggaatcccttcaaaagatttcttcggga 1200
QY 1201 TTACCGGATCGCTTCAAGTTCGATGCTGATGTTGGATGTTGGGAAAGCCTTGCA 1260
DB 1201 ttaccggtcgcttaagttcgattcgatggtgtagattttggatggggaagcctgcaa 1260
QY 1261 AATTGACATACCTCTGTTGATTATGCAGAAATGATTATGTTATGATTCAGTCCAGGAT 1320
DB 1261 aattgacattacctctgttgattatgcagaattgatttatgtgatccagtcaggggatt 1320
QY 1321 TTGAAAAGGTGTGGAGATTGGAGTATCATTCCTTAAGATTTCATATGATGATTCAT 1380
DB 1321 ttgaaaaggtgtgagattgagatgatacttgccatgattcattatgattgcaatg 1380
QY 1381 AAATCTTTGAAGAGGCTTTTGTCTTTGTCATAGTCTCTTTAATAGAACCATATTTGCT 1440
DB 1381 aaatctttgaagaggcttttgcctttgtcatagctcttctaataagaacacatttgc 1440
QY 1441 GCAATAAGTACCAAGTCCCTTTAGTAACTACACCAACCCCTACTTTCGAGCGGGAAC 1500
DB 1441 gcaataagtaccagtcctttagtaacactacacccaacctacttccgagggcgaac 1500
QY 1501 ACCACAACGAGGTTCAATCACTAGAAAGTTGTACTTCAATAATTCAGAGGTCGAATATA 1560
DB 1501 accacaacgaggttcaatcactagaagttgtacttcataaattccagagggtcgaaata 1560
QY 1561 CACGGTTGCTCTGAAAGTTGAACCTCACCTGACATGCTGATGGTGTACGATAGTATTGT 1620
DB 1561 cacggttgcctctgaaagttgaacctcacacctgacatggtgtttacgataggtattgt 1620
QY 1621 ATATGCAATATATATCTTCCATAAAGTATCCCTATGCAATAGAGAAACATGTTATGCTTA 1680
DB 1621 atatgcaattatatacttccataaagtatacctatgcaatagagaaacatgttattgtta 1680
QY 1681 AAAAAAAAAAAAAAAAAAAAAA 1703
DB 1681 aaaaaaaaaaaaaaaaaaaaaa 1703

Db 690 taccgacgacgcgcatcgaaccacgcttcgttttccaccaatcccaaaattaaagaaatg 749
QY 795 AAGAACAAAGTACTGAATCTCAGAGATCCGAAACCCACAAATAGCTGTAAAGACGTTTACA 854
Db 750 aag-----ggttgattcagtcagagattccaagtttagtccatctctcatctttgt 803
QY 855 ATGACGTGGATACGATATGGACATGCATGTCATCAATCAAAAGATGACGTCGTATCAGAG 914
Db 804 gcatgacgcttatatgtgggctggcataaagaaatcattccacagcatgaag----- 858
QY 915 GAATCATGAACGACGAAATAGCTCGAGTACTTTCAGTTCATACACGGGATTCGCCGAGGA 974
Db 859 -----accagacacagcagagatgcatctttctgattccgctcgatctaaagcca 908
QY 975 CTCTCAGCCCGCTGTCGGCTACTACTTTGGCAACTGCTTCGGTCATCGCTTGCA 1034
Db 909 cdatagatcccgctgctctgaaatctctcggaactgcttaccgacgctgccc 968
QY 1035 AAAGCAACACATAAAGAGTTAGTTGGGGATTAAGGGCTTCTTCTTCAGTTGCAGCTATT 1094
Db 969 agaatcgccgagagctggtggagagaaagggtttctgagcgtcaggtgaatc 1028
QY 1095 GGAGAGCCATTGAAAGAGGTTCACACAGAAAGAGCGTTCCTGCAGATGCAAAACT 1154
Db 1029 gcgccgagataaaaaaaggat---caacgacaaagagaattattagaacggtggagaa 1085
QY 1155 TGGTTATCGAATCTAATGGAATCCCTTCAAAAGATTTCTCGGGATTTACCGGATCCCT 1214
Db 1086 tggtcgagagatcgtgaagcgttcgagaaatcatattttcgtgagcagtcgagc 1145
QY 1215 AAGTTCGATTCGTATGCTGATGATTTGGATGGGAAAGCCTGCAAAATTTGACATTACC 1274
Db 1146 aagctagatcttcaaggtcagatcttgatg999gaagcgagaaagcaagaaatcattg 1205
QY 1275 TCTGTTGATTATCAGAAATG---ATTATGATTCAGTCCAGGATTTTGAAGAAAGGT 1331
Db 1206 tcgattgagggagaaatgatgcaatgacrcrttctgaaagccaggatctcgaaggaga 1265
QY 1332 GTGAGATTTGGAGTATCATTCGCTAAGATTCATATGATGCATTTGCAAAAATCTTTTGA 1391
Db 1266 ttggaggttctgtcttcttgcttaaggacaaatggtctgttctgtcttattttca 1325
QY 1392 GAAGGCTTTTGTCTTGTTCATAGTCTCTTTAATAGAACCATATTTGCTGCAATAAGTA 1451
Db 1326 ctg99gaattaaaggtttaaataatgctgtaattaaactaatatctattgtacaactaa 1385
QY 1452 CCAAGTCTTTTGTAGTAACACTACACCAAAACCCCTACTTT 1488
Db 1386 ttaagtgtagtaacgtggaagaataatccctattat 1422

RESULT 6
AAF58252/c
ID AAF58252 standard; DNA; 936 BP.
XX
AC AAF58252;
XX
DT 24-APR-2001 (first entry)
XX
DE Oligonucleotide D1835.
XX
KW Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
OS Synthetic.
XX
PN W0200107665-A2.
XX
PD 01-FEB-2001.
XX
PF 26-JUL-2000; 2000HO-US20476.
XX

PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX Umek RM;
XX WPI; 2001-159728/16.
DR
PT Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 127; 159pp; English.
PS
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
SQ Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 6.5%; Score 110; DB 22; Length 936;
Best Local Similarity 1.1%; pred. No. 8.7e-20;
Matches 8; Conservative 434; Mismatches 264; Indels 0; Gaps 0;

QY 998 TAACTACTTTGGCAACTGCTTCGCTCATGCTTGGCAAAAGCAACATAAAGAGTTAGT 1057
Db 762 WWWWWWXX 703
QY 1058 TGGGGATAAAGGGCTTCTTGTTCAGCTTGCAGCTATTGGAGAAGCCATTGAAAAGAGTT 1117
Db 702 WWWWWWXX 643
QY 1118 GCACAAGCAAAAGCGCTTCTTCAGATGCAAAAGCTTGGTATCGAATCTAATGAAT 1177
Db 642 WWWWWWXX 583
QY 1178 CCCTTCAAAAGATTTCTCGGATTAACGGATCGCCCTAAGTTCGATTCGTATGCTAGA 1237
Db 582 WWWWWWXX 523
QY 1238 TTTTGGATGGGAAAGCCCTCAAAATTTGACATTTACCTCTGTTGATTATCCAGAAATGAT 1297
Db 522 WWWWWWXX 463
QY 1298 TTATGTGATTTCAGTCCAGGATTTTGAAGAAAGCTGTGGAGATTGGAGTATCATTCGCTAA 1357
Db 462 WWWWWWXX 403
QY 1358 GATTCATATGGATGCATTTTGCAGAAATCTTTGAAGAGGCTTTTGTCTTTGTATAGTC 1417
Db 402 WWWWWWXX 343
QY 1418 TCTTTAATAGAACCATATTTGCTGCAATAAAGTACCAAGCTTTTAGTACACATACACCA 1477
Db 342 WWWWWWXX 283
QY 1478 AACCTTACTTTTCGAGCGGGAACACACAGAGGTTTCAATCACTAGAAGGTTGTACTTC 1537
Db 282 WWWWWWXX 223
QY 1538 ATAAATTCAGAGGTGCAATATACACCGTTTGTCTCTCTGAAAGCTTGAACCTCACACCTGA 1597
Db 222 WWWWWWXX 163
QY 1598 CATGGTGTACGATAGGATTTGTATTAATGCCATTATATACTTCCATAAAGATTCCTATGC 1657
Db 162 WWWWWWXX 103

Db 517 576
Qy 1538 ATAAATCCAGAGTCGAATATACACCGTTCTCTCGTAAAGTTGAACCTTCACACCTGA 1597
Db 636
Qy 1598 CATGGTGTACGATAGGTATTTGTATAATGCCATTATATCTTCCATTAAGTATCCTATGC 1657
Db 696
Qy 1658 AATAGAGAACATGTATGTGTTAAAAA..... 1703
Db 742

RESULT 13
AAF58254
ID AAF58254 standard; DNA; 936 BP.
AC AAF58254;
XX
XX 24-APR-2001 (first entry)
DT
DE Oligonucleotide D1875.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
OS
XX WO200107665-A2.
PN
XX 01-FEB-2001.
PD
XX 26-JUL-2000; 2000WO-US20476.
PF
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
PI Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface -
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
XX
XX Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;
SQ

Query Match 6.3%; Score 106.8; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 6.3e-19;
Matches 6; Conservative 434; Mismatches 266; Indels 0; Gaps 0;

Qy 998 TAACTACTTTGGCAACTGTCTTGGTCATCGGTTGCAAAAGCAACATAAAGATTAGT 1057
Db 96
Qy 1058 TGGGATAAAGGGCTTCTGTTCAGTTGCAGCTATTGGAGAGCCATTGAAAGAGGTT 1117
Db 156

Qy 1118 GCACAAGAAAAAGCGCTTCTTCGAGATGCAAAAACCTTGGTTATCGGAATCTAATGAAT 1177
Db 216
Qy 1178 CCCTTCAAAAGATTCTTCGGGATTACCGGATTCGCCCTAAGTTCGATTCGATGCTGATGA 1237
Db 276
Qy 1238 TTTTGGATGGGAAAGCCTGCAAAATTTGACATTACCTCTCTGTTGATTATGCAGAAATTGAT 1297
Db 336
Qy 1298 TTATGTGATTTCAGTCCAGGATTTTGAAGAGGTGTGAGATTGGAGTATCATTCATTCCTAA 1357
Db 396
Qy 1358 GATTCATATGATGATGATTTGCAAAAATCTTTGAAGAGGCTTTTGTCTTTGTCTATAGTC 1417
Db 456
Qy 1418 TCTTTAATAGAACCATATTTGCTGCAATAAAGTACCAAGTCTCTTTAGTAACACTACACCA 1477
Db 516
Qy 1478 AACCTACTTTTCGAGCGGGAACACACACAGAGTTTCACTACACTAGAAGGTTGACTTC 1537
Db 576
Qy 1538 ATAAATCCAGAGTCGAATATACACCGTTGCTCTGTAAGAGTTGAACCTTCACACCTGA 1597
Db 636
Qy 1598 CATGGTGTACGATAGTATTTGTATAATGCCATTATATCTTCCATTAAGTATCCTATGC 1657
Db 696
Qy 1658 AATAGAGAACATGTATGTGTTAAAAA..... 1703
Db 742

RESULT 14
AAF58257
ID AAF58257 standard; DNA; 936 BP.
AC AAF58257;
XX
XX 24-APR-2001 (first entry)
DT
DE Oligonucleotide D1954.
XX
XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
OS
XX WO200107665-A2.
PN
XX 01-FEB-2001.
PD
XX 26-JUL-2000; 2000WO-US20476.
PF
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:04:15 ; Search time 168.74 seconds
(without alignments)
1800.666 Million cell updates/sec

Title: US-08-894-356C-3
Perfect score: 1605
Sequence: 1 TGTCTGACGAATCCATTCA.....TTAAAAA.....1605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	67	4.2	7218	1 US-08-232-463-14	Sequence 14, Appl
C 2	45	2.8	2852	3 US-09-027-137-2	Sequence 2, Appl1
C 3	44.8	2.8	1474	4 US-08-821-994-64	Sequence 64, Appl
C 4	43.8	2.7	3157	6 5198347-3	Patent No. 5198347
C 5	43.6	2.7	1582	3 US-08-545-196B-10	Sequence 10, Appl
C 6	43.6	2.7	1582	3 US-08-545-196B-12	Sequence 12, Appl
C 7	43.6	2.7	1882	4 US-09-370-253-1	Sequence 1, Appl1
C 8	43.4	2.7	1046	1 US-08-361-467B-4	Sequence 4, Appl1
C 9	43.4	2.7	1332	2 US-08-484-332C-4	Sequence 4, Appl1
C 10	43.2	2.7	1332	2 US-09-057-762-1	Sequence 1, Appl1
C 11	43.2	2.7	1332	3 US-08-326-119A-1	Sequence 1, Appl1
C 12	41.8	2.6	3337	1 US-08-072-610-1	Sequence 1, Appl1
C 13	41.8	2.6	3337	2 US-08-719-822B-1	Sequence 1, Appl1
C 14	41.8	2.6	3337	4 US-09-092-458-1	Sequence 1, Appl1
C 15	41.6	2.6	581	2 US-08-557-309B-22	Sequence 22, Appl
C 16	41.6	2.6	581	3 US-08-834-306-22	Sequence 22, Appl
C 17	41.6	2.6	581	4 US-08-993-674A-22	Sequence 22, Appl
C 18	41.6	2.6	5852	1 US-07-867-106-2	Sequence 2, Appl1
C 19	40.8	2.5	746	4 US-09-013-810-1	Sequence 1, Appl1
C 20	40.6	2.5	240	1 US-08-628-417-6	Sequence 6, Appl1
C 21	40.4	2.5	3527	2 US-08-909-965C-7	Sequence 7, Appl1
C 22	40	2.5	140	1 US-08-628-417-5	Sequence 5, Appl1
C 23	40	2.5	1798	2 US-08-557-128-12	Sequence 12, Appl
C 24	40	2.5	2295	1 US-08-375-300-3	Sequence 3, Appl1
C 25	40	2.5	2295	3 US-09-177-431-3	Sequence 3, Appl1
C 26	40	2.5	2295	5 PCT-US95-16930-3	Sequence 3, Appl1
C 27	40	2.5	4080	1 US-08-375-300-1	Sequence 1, Appl1

28	40	2.5	4080	3 US-09-177-431-1	Sequence 1, Appl1
29	40	2.5	4080	5 PCT-US95-16930-1	Sequence 1, Appl1
30	39.8	2.5	358	2 US-08-721-488-9	Sequence 9, Appl1
31	39.4	2.5	923	4 US-09-004-731-66	Sequence 66, Appl1
32	39.4	2.5	923	4 US-08-749-699-66	Sequence 66, Appl1
C 33	39	2.4	1736	3 US-09-182-816-22	Sequence 22, Appl1
C 34	39	2.4	1736	3 US-09-182-816-24	Sequence 24, Appl1
C 35	39	2.4	1736	3 US-09-471-528-22	Sequence 22, Appl1
C 36	39	2.4	1736	3 US-09-471-528-24	Sequence 24, Appl1
C 37	38.8	2.4	8920	2 US-08-446-855A-1	Sequence 1, Appl1
C 38	38.8	2.4	8920	4 US-09-150-741-1	Sequence 1, Appl1
C 39	38.6	2.4	6877	3 US-08-347-340-1	Sequence 1, Appl1
C 40	38.4	2.4	260	2 US-08-520-678A-29	Sequence 29, Appl1
C 41	38.4	2.4	2230	4 US-08-378-313-24	Sequence 24, Appl1
C 42	38.2	2.4	923	3 US-08-906-769-144	Sequence 144, App
C 43	38.2	2.4	923	3 US-08-906-616-144	Sequence 144, App
C 44	38.2	2.4	923	3 US-08-639-075A-144	Sequence 144, App
C 45	38.2	2.4	923	4 US-09-012-431-144	Sequence 144, App

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFELINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Larøner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22133-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DCS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9pt-Fls
US-08-232-463-14

```

Query Match      4.2%; Score 67; DB 1; Length 7218;
Best Local Similarity 2.0%; Pred. No. 2.7e-08;
Matches 7; Conservative 221; Mismatches 121; Indels 0; Gaps 0;

Qy 1044 TTGTACTACCAAGCCGATCGAGTTCCTGGTGGGATGATACAAAGCGATCGTGA 1103
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 1442 TTGTACTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1383
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1104 GCATCACCTAGGCAATGTAAGAAACAAAGAGTGGAGAGCAACCCGAGATCTT 1163
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1164 TCAGTACAAAGATCGTGAGTCAACTGTGCTGTGGAGTTCGCCAAGGTTCAAGGT 1223
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1224 TTACCACGTGATTTGATGGGGAACCCAGAGTGTGAGGAGTGTTCGAACAATAG 1283
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1284 GTTGTGATGAATGGTGTATTGTACCAAGCAAAATGGAGGAAGCATTTGATGTGA 1343
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1344 GATTAGTTGGAACAAATGCTATGGAGAGTGTGAGAAAGATAAAGAG 1392
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1094
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 2

```

US-09-027-137-2/c
; Sequence 2, Application US/09027137
; Patent No. 6013450
; GENERAL INFORMATION:
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Yue, Henry
; TITLE OF INVENTION: CAF1-RELATED PROTEIN
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Dr.
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/027,137
; APPLICATION NUMBER: 59
; FILING DATE: Filed Herewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0476 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-855-0555
; TELEFAX: 650-845-4166
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2852 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: PROSNOT16

```

```

; CLONE: 2229466
US-09-027-137-2

Query Match      2.8%; Score 45; DB 3; Length 2852;
Best Local Similarity 56.4%; Pred. No. 0.014;
Matches 84; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

Qy 1453 GAGCTGCTATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1512
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 2599 GAGCGGCCCCCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2540
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1513 TTGTTTCTTATGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1572
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 2539 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 2480
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1573 ATGTTTCTAGTTTATTAATAAAAAAAAAAAAAA 1601
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 2479 TTTATTTTACAAAATTAATAATAATAATA 2451

RESULT 3
US-08-821-994-64/c
; Sequence 64, Application US/08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier Rp
; APPLICANT: Jepson, Ian
; TITLE OF INVENTION: Promoters
; FILE REFERENCE: PPD 50108
; CURRENT APPLICATION NUMBER: US/08/821,994A
; CURRENT FILING DATE: 1997-03-22
; EARLIER APPLICATION NUMBER: PCT/GB97/00729
; EARLIER FILING DATE: 1997-03-18
; EARLIER APPLICATION NUMBER: GB 9606062.9
; EARLIER FILING DATE: 1996-03-22
; NUMBER OF SEQ ID NOS: 89
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 64
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Brassica napus
US-08-821-994-64

Query Match      2.8%; Score 44.8; DB 4; Length 1474;
Best Local Similarity 59.4%; Pred. No. 0.011;
Matches 76; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

Qy 1475 TGATGTTTATTTTCTTTTGGAGGGTCTTCTTTTCTTTTCTTTTCTTTTCTTT 1534
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 1448 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1389
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1535 CTCTACGTTATGAGAGAAACCCGAGTATAAGGAATAATGTTTCTAGTTATTA 1594
      |||| ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 1388 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1329
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Qy 1595 AAAAAAAA 1602
      |||| |||
Db 1328 CAAATCAA 1321

RESULT 4
5198347-3
; Patent No. 5198347
; APPLICANT: Miller, Louis H.; ADAMS, JOHN H.; KASLOW,
; DAVIC C.; FANG, XIANGDOUG
; TITLE OF INVENTION: DNA ENCODING PLASMODIUM VIVAX AND
; PLASMODIUM KNOWLESII DUFFY RECEPTOR
; NUMBER OF SEQUENCES: 27
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/554,837

```

[illegible]

APPLICANT: Allen,

APPLICANT: Allen,

; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370.253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Hordeum vulgare
US-09-370-253-1

Query Match 2.7%; Score 43.6; DB 4; Length 1882;
Best Local Similarity 57.2%; Pred. No. 0.027;
Matches 79; Conservative 0; Mismatches 59; Indels 0; Gaps 0;
Qy 1444 TTATTTATGAGCTGCTATGACTCACATGCATGATGTTATTTTGGAGGGTTCT 1503
Db 1879 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1820
Qy 1504 TTCTTTTATTTGTTCTATGTTTCTTCTGCTGCTATGAGAGAAACCGAGTAT 1563
Db 1819 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 1760
Qy 1564 AAGAGATAATGTTTCA 1581
Db 1759 TCGGAATAATCTATGCA 1742

RESULT 8
US-08-361-467B-4/c
; Sequence 4, Application US/08361467B
; Patent No. 5633441
; GENERAL INFORMATION:
; APPLICANT: De Greef, Willy
; APPLICANT: Van Emmelo, John
; APPLICANT: De Oliveria, Dulce E.
; APPLICANT: De Souza, Maria-Helena
; APPLICANT: Van Montagu, Marc
; TITLE OF INVENTION: PLANTS WITH MODIFIED FLOWERS, SEEDS OR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/361.467B
; FILING DATE: 22-DEC-1994
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681.492
; FILING DATE: 04-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP90/01275
; FILING DATE: 01-AUG-1990
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 89 402 224.3
; FILING DATE: 04-AUG-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Schulman, Robert M.

; REGISTRATION NUMBER: 31,196
; REFERENCE/DOCKET NUMBER: 010830-027
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1046 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; IMMEDIATE SOURCE:
; CLONE: 3C9
US-08-361-467B-4

Query Match 2.7%; Score 43.4; DB 1; Length 1046;
Best Local Similarity 57.9%; Pred. No. 0.023;
Matches 77; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 1473 CATGTATGTTTATTTTATTTTGGAGGGTCTTCCCTTTATGTTTCTATGTTTCT 1532
Db 1042 CTGCAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 983
Qy 1533 TTCTGTAGCTTATGAAGAGAAACCGAGTATAAGGAATAATGTTTTCAGTTATTAAAA 1592
Db 982 TTTTATTTTATTTTATTTTGAAGAAACTTATAATCCATTTACTTGTGGTGTGATA 923
Qy 1593 AAAAAAATAAAAA 1605
Db 922 GAGAGAAACAAAA 910

RESULT 9
US-08-484-332C-4/c
; Sequence 4, Application US/08484332C
; Patent No. 5767374
; GENERAL INFORMATION:
; APPLICANT: De Greef, Willy
; APPLICANT: Van Emmelo, John
; APPLICANT: De Oliveria, Dulce E.
; APPLICANT: De Souza, Maria-Helena
; APPLICANT: Van Montagu, Marc
; TITLE OF INVENTION: PLANTS WITH MODIFIED FLOWERS, SEEDS OR
; NUMBER OF SEQUENCES: 13
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: BURNS, DOANE, SWECKER & MATHIS
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/484.332C
; FILING DATE: 7-JUNE-1995
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/361.467
; FILING DATE: 22-DEC-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/681.492
; FILING DATE: 04-APR-1991
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: WO PCT/EP90/01275
; FILING DATE: 01-AUG-1990
; PRIOR APPLICATION DATA:

[illegible]

```

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999

```

	Query Match	2.6%	Score 41.8;	DB 2;	Length 3337;
	Best Local Similarity	51.3%;	Pred. No. 0.11;	Mismatches 0;	Gaps 0;
	Matches 97;	Conservative		Indels 0;	
QY	227	AAGTGCTGAAAAACTTTGCACGACGCGTGGAAAAATTAAGAATGGACTGCCCTTAGTAT	286		
Db	2167	AAGAAGTGAAGAGGTACCCCGACGAAGTGAAGAAGTGAAGAAGTACCAAGAAAGATTGG	2226		
QY	287	TGGTGGATTTCATACTAGCTGGGAAACTTGGAAAAGATGAAGAAGGGGTTTTTCAGGG	346		

Db 2227 AACAGGTACCAGAGAGTGAAGAGGTACCAGAGAGTGAAGAGGTACCAGAGAG 2286
QY 347 TGAATACGACGATGATGGTGTAGAGGTGACAGTGGCTGTTCAGAGAGATAG 406
Db 2287 TGAAGAAGTGAAGAAGTAGAAGAAGTAGAGGTACCAGCGGTAGTAGAAGTAC 2346
QY 407 AAGTTGCAG 415
Db 2347 CAGCGGTAG 2355

RESULT 14
US-09-092-458-1
; Sequence 1, Application US/09092458
; Patent No. 6231861
; GENERAL INFORMATION:
; APPLICANT: Barnwell, John
; TITLE OF INVENTION: Plasmodium vivax Blood Stage Antigens,
; TITLE OF INVENTION: Monoclonal Antibodies, and Diagnostic Assays
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Darby and Darby
; STREET: 805 Third Ave.
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10022-7513
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/092,458
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/719,821
; FILING DATE: 09/30/96
; ATTORNEY/AGENT INFORMATION:
; NAME: Gogoris, Adda
; REGISTRATION NUMBER: 29,714
; REFERENCE/DOCKET NUMBER: 5986/17686US3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212)527-7700
; TELEFAX: (212)753-6237
; TELEX: 236687
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3337 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Plasmodium vivax
; IMMEDIATE SOURCE:
; CLONE: pVMB3.3.1
US-09-092-458-1

Query Match 2.6%; Score 41.8; DB 4; Length 3337;
Best Local Similarity 51.3%; Pred. No. 0.11;
Matches 97; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 227 AAGTGCTGAAACTTTCAGCAGCGGTGAGAAATTAAGATGGACTGCCTTAGTAT 286
Db 2167 AAGAAGTAGAAGAGGTACCCGAGAGAGTAGAAGAGTGAAGAGGTACCAAGAGATGG 2226
QY 287 TGGTGGATTCTCACTAGCTGGGAAACTTGGAAAGATGAAGAGGGGTTTCAGGG 346

Db 2227 AACAGGTACCAGAGAGTGAAGAGGTACCAGAGAGTGAAGAGGTACCAGAGAG 2286
QY 347 TGAATACGACGATGATGGTGTAGAGGTGACAGTGGCTGTTCAGAGAGATAG 406
Db 2287 TGAAGAAGTGAAGAAGTAGAAGAAGTAGAGGTACCAGCGGTAGTAGAAGTAC 2346
QY 407 AAGTTGCAG 415
Db 2347 CAGCGGTAG 2355

RESULT 15
US-08-557-309B-22/c
; Sequence 22, Application US/08557309B
; Patent No. 5916572
; GENERAL INFORMATION:
; APPLICANT: Reed, Steven G.
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Lodes, Michael J.
; APPLICANT: Houghton, Raymond L.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THE DETECTION AND
; PREVENTION OF
; NUMBER OF SEQUENCES: 69
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: SEED and BERRY LLP
; STREET: 6300 Columbia Center, 701 Fifth Avenue
; CITY: Seattle
; STATE: Washington
; COUNTRY: USA
; ZIP: 98104-7092
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/557,309B
; FILING DATE: 14-NOV-1995
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Maki, David J.
; REGISTRATION NUMBER: 31,392
; REFERENCE/DOCKET NUMBER: 210121.422
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (206) 622-4900
; TELEFAX: (206) 682-6031
; INFORMATION FOR SEQ ID NO: 22:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 581 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-557-309B-22

Query Match 2.6%; Score 41.6; DB 2; Length 581;
Best Local Similarity 64.6%; Pred. No. 0.052;
Matches 62; Conservative 0; Mismatches 34; Indels 0; Gaps 0;
QY 1510 TTATTGTTTTCTATGTTTTTTCTTTCTTTGATGAGAGAAACCGAGTATAAGGA 1569
Db 581 TTGAACAAAAACACACAA 522
QY 1570 ATATGTTTTTCAGTTATTAACAAAAACAAAAA 1605
Db 521 GACATCTCTCAATAAAAAACAGACATATAAAAA 486

Search completed: November 5, 2001, 18:04:28
Job time: 15068 sec

7

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:43 ; Search time 5816.79 Seconds
(without alignments)
2608.282 Million cell updates/sec

Title: US-08-894-356C-3
Perfect score: 1605
Sequence: 1 TGTCCGAGGAATCCATTCA.....TTAAAAAAAAAAAAAAAAAAAA 1605

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*

44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estrol:*
83: em_estrol2:*
84: em_estrol3:*
85: em_estrol4:*
86: em_estrol5:*
87: em_estrol6:*
88: em_estrol7:*
89: em_estrol8:*
90: em_estrol9:*
91: em_estrol10:*
92: em_estrol11:*
93: em_estrol12:*
94: em_estrol13:*
95: em_estrol14:*
96: em_estrol15:*
97: em_estrol16:*
98: em_estrol17:*
99: em_estrol18:*
100: em_estrc19:*
101: em_estrc20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				TITLE	
Result No.	Score	Query Match	Length DB ID	JOURNAL COMMENT	Generation of ESTs from potato leaves and petioles
-----				Unpublished (2000):	
-----				Contact: Cathy Ronning	
-----				The Institute for Genomic Research	
-----				For clone request: please contact Research Genetics, Libraries	
-----				Division tel 1-800-711-6195, email cdna@resgen.com.	
-----				Location/Qualifiers	
-----				1. 632	
-----				/organism="Solanum tuberosum"	
-----				/cultivar="Kennebec"	
-----				/db_xref="taxon:4113"	
-----				/clone="cSTB2C8"	
-----				/clone_lib="potato leaves and petioles"	
-----				/tissue_type="leaflets and petioles"	
-----				/dev_stage="8 weeks old plants"	
-----				/lab_host="SOLR"	
-----				/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Tissue was supplied by Dr. Fry (Cornell University). Leaflets and petioles were isolated from 8 week old greenhouse grown plants. The plants were watered and fertilized freely. The tissue was immediately frozen in liquid nitrogen."	
-----				BASE COUNT 183 a 181 c 95 g 173 t	
-----				ORIGIN	
-----				Query Match 26.7%; Score 428; DB 141; Length 632;	
-----				Best Local Similarity 81.2%; Pred. No. 3.6e-95;	
-----				Matches 497; Conservative 0; Mismatches 115; Indels 0; Gaps 0;	
QY	843	CTCAGGAGAGCGCCATTCTCCACATTTCCAATCTCTCCGACACGCTGGCTAGCCGT	902		
DB	631	CTCAATAGCCGCTGCTCTCTACITTCAGTCCCTCTCCGACACGCTGGCTAGCTGT	572		
QY	903	CACAGTGGCGGCCAACTCAACCGGAGACTACACTGTGTACACTGTGTTCGTGATTTG	962		
DB	571	CACAGTGGCGGCCAACTCAACCGGAGACTACACTGTGTACACTGTGTTCGTGATTTG	512		
QY	963	CAGGAAAGGGTTGATCTCTCAATGCCAGAAAGTTACTTCGCAACCTAAATTCAGCG	1022		
DB	511	CGGAAAGAGTGGATCTCTCAATGCCAGAAAGTTACTTCGCAACCTAAATTCAGCG	452		
QY	1023	TTTTCAGTGGCGGCCAGTCTTTTGTACTAGCAAGCCGATCGAGTTCGCTGGTGGAT	1082		
DB	451	TTTTCAGTGGCGGCCAGTCTTTGCTGTATCGATCCGATTTGCGCGAGGAT	392		
QY	1083	GATACAAAGCGATCGTGAACGATGACCTAAGGCCATTGATGAAGAAACAGAGATG	1142		
DB	391	GATTGACGAAGCAATTTGTAATCAGATCGAAGCGATTGATGAAGGAACAGAGATG	332		
QY	1143	GGAGAGCAACCGAAGATCTTTTCACTACAAAGATGCTGGAGTGAATTTGCTGTTGG	1202		
DB	331	GGAAAGTAGTCCGAGATATTTGAGTACAAAGATGCCGAGTGAATTTGCTGGGTGG	272		
QY	1203	AAGTTCCGCAAGTTTCAAGTTTACGATTTTGGATTTTGGATGGGAAACCCAGAGATG	1262		
DB	271	AAGTTCCGCGAGTTTAAAGTGTATGATGTGATTTGGATGGGAAACCCAGAGATG	212		
QY	1263	GAGGAGTGGTTCGACAAATAGTTTGTATGGAATTTGATGATTTGATCAAGGCAAAATG	1322		
DB	211	GAGAATGTTGTAACAATAGTTTGTATGGAATTTGATGGAATTTGATCAAGGCAAAATG	152		
QY	1323	AGGAAGCAATTTGATGCGAGATTTGGAACCAATGCTATGAGAGGTTGGAGAA	1382		
DB	151	AGGAAGCAATTTGATGCGAGATTTGGAACCAATGCTATGAGAGGTTGGAGAA	92		
QY	1383	AGATAAAGATTCCTCATCGAAACTGCTTAATTTGCTTGGACTCAACTGGCTACA	1442		
DB	91	GGATAAAGATTTCTTATGGAAGACTTAATTAATTTGCTTTGCTTAATTTGATGTT	32		
QY	1443	CTTTATTTATGA	1454		
DB	31	TTGTTTTTTTAA	20		

BE19635 EST423404

AW649974 EST328428

BG126859 EST472505

BF051093 EST436268

BE19634 EST423403

AW928744 EST337532

BG129408 EST475054

AW618433 EST320419

BE920362 EST424131

AW039511 EST281792

AW649611 EST328065

BF053244 EST438474

BF275133 GA_EB002

AW041172 EST284036

AW617090 EST323501

BE458676 EST413968

BG096122 EST450641

BE924264 EST428033

BE353229 EST400361

BG446054 GA_Ea003

AW731285 GA_Ea003

BF054047 EST439277

AW639481 EST309981

AW668525 GA_Ea001

AW617259 EST323670

BG441719 GA_Ea001

AW728139 GA_Ea001

AW189979 EST2404.Y

AW735992 EST336760

BG441701 GA_Ea001

AW130615 BNLGH1741

AW617396 EST323807

BF053972 EST439202

AW172325 BNLGH1775

AW616992 EST323403

AW1730339 BNLGH1668

AW690825 NF035D15

AW399665 EST310165

AW1901068 sc19b1.Y

AW1730449 BNLGH1720

BF268244 GA_EB000

AW1730451 BNLGH1721

BF278351 GA_EB003

AW934606 EST353498

AW1730368 BNLGH1675

BE19635

EST423404

ESTB2C8 3' sequence, mRNA

BE19635

EST.

BE19635.1

GI:10445711

potato.

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 632)

van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning ,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.

BE19635

EST423404

ESTB2C8 3' sequence, mRNA

BE19635

EST.

BE19635.1

GI:10445711

potato.

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 632)

van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning ,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.

BE19635

EST423404

ESTB2C8 3' sequence, mRNA

BE19635

EST.

BE19635.1

GI:10445711

potato.

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 632)

van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J., Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning ,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.

```
RESULT 2
AW649974 LOCUS AW649974 657 bp mRNA EST 04-APR-2000
DEFINITION EST328428 tomato germinating seedlings, TAMU Lycopersicon
esculentum cDNA clone cLE11l6 5', mRNA sequence.
ACCESSION AW649974
VERSION AW649974.1 GI:7411212
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 657)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S, Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley
,S.D.
TITLE Generation of ESTs from germinating tomato seed
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
source Location/Qualifiers
1..657
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLE11l6"
/tissue_type="whole seedlings"
/dev_stage="7 days post imbibition"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; 7 days post imbibition on water-agar. Mixed stage
whole germinating seedlings from seed coat emergence up
to two centimeters in seeds not showing obvious signs of
germination were discarded."
BASE COUNT 212 a 120 c 147 g 178 t
ORIGIN

Query Match 26.2%; Score 420.2; DB 119; Length 657;
Best Local Similarity 80.7%; Pred. No. 3.le-93;
Matches 517; Conservative 0; Mismatches 118; Indels 6; Gaps 2;

Qy 96 TACAAAGTGAAGTCTCGAAGAAACAAACGCTGAACACCATAAACACTAGGAAAAA 155
Db 17 TACAAATGTGAACATCTTGAAGAAATCAATGTGAAGCCAAACCACTAGGAAAAA 76

Qy 156 AGAGTGCAATTGTGTAACATTTGATCTCTCTTACCTAGCTTCTATTACACCAAAATT 215
Db 77 AGAATGTCAATTAGTACATTTGATTTACCTTACTTACTTATTTATTATACCAAAAGTT 136

Qy 216 TCTCATCTATAA---AGGTGCTGAAACCTTTTGACGAGACGCTGGAAAAAATTAAGATGG 272
Db 137 GATGGTTTACAAAGTTAGGGCTGAAGCTTTGAGAAACAGTAGAAGAAATTTGAAGATGG 196

Qy 273 ACTGGCCTTAGTATTTGGTGGATTCTTATCAACTAGCTGGGAAACTTGAAGATGAAGA 332
Db 197 TTTGGCTTTAGTGTGGAAGATTTTATCACTTTGCTGGAAAAATTTGGAAAAAGATGA 256

Qy 333 AGGGGTTTTACGGTGAATACGACGATGACATGGTGTCTAGAGGTGACAGTGGCTGT 392
Db 257 AGGAGTATTTAAGTGGAAATATGATGATGATGATGATGATGATGATGATGATGATG 316
```

```
393 TGCAGAGAGATAGAGTTCGACATCTTACTG---ATGAAGAAGGCACCAAAATTGCA 449
Db 317 GGCTCAAGAAATTCAGTTGCTAATCTTACTGATCATCATGAAGGATCAATAAATCCA 376

Qy 450 GGACTTGAATTCCTTGAATAAAATCTTGAATTTGGAAGGGCTTCATCGCCTCTTCTAGC 509
Db 377 AGATTGTGATACCTTTATAACAAAATCTTGAATTTAGAAGGGCTTCATAGGCCTCTACTCGC 436

Qy 510 TGTGCACTCACCAGAGCTCAAGGACGGCTCACCATGGGATTAGCATTTAACCATGCTGT 569
Db 437 CGTACAGCTCACCAGAGCTTAAAGACGGCTGGCAATGGATTAGCATTTAATCATGCTGT 496

Qy 570 GCTGGATGGTACTTCGACGTGGCACCTTTATGACCTCGTGGTCCGAGCTTTGCTGTGGGTC 629
Db 497 GCTGGAGGTACTTCGACGTGGCACCTTCATGACTTCTGGCCCGACGCTTGTAGTGGGGC 556

Qy 630 CACCTCAATTTCTGTCACCATTCCTTTGAACGAACCAAGGCTCGTACACTCGAGTCAA 689
Db 557 CACCTCCATCTCGGTCCGACCTTTCTCGAAAGAACTAAAGCTCGTGACACTCGAGTCAA 616

Qy 690 GCTCAACCTCTCTCAACCATCAGATGCACCGCAACATGCTA 730
Db 617 ACTCAACCTATCTAAGCCATCCGACGCGCGGAACATGCCA 657

RESULT 3
BG126859 LOCUS BG126859 649 bp mRNA EST 31-JAN-2001
DEFINITION EST472505 tomato shoot/meristem Lycopersicon esculentum cDNA clone
cTOF13L13 5' sequence, mRNA sequence.
ACCESSION BG126859
VERSION BG126859.1 GI:12627047
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 649)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Uterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
TITLE Generation of ESTs from tomato shoot/meristem tissue
JOURNAL Unpublished (2001)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
FEATURES
source Location/Qualifiers
1..649
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cTOF13L13"
/tissue_type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
BASE COUNT 177 a 170 c 148 g 154 t
ORIGIN

Query Match 25.3%; Score 406.4; DB 174; Length 649;
Best Local Similarity 79.3%; Pred. No. 7.8e-90;
Matches 514; Conservative 0; Mismatches 116; Indels 18; Gaps 2;

Qy 385 GTGCTGTTGCAAGAGAGATAGAAGTTGCAGATCTTACTG---ATGAAGAAGGCACCAACC 441
Db 385 GTGCTGTTGCAAGAGAGATAGAAGTTGCAGATCTTACTG---ATGAAGAAGGCACCAACC 441
```


Db 2 GTGGATGAGGCTCAGAAAATTCAGTTGCTAATCTTACTGATCATCATGAAGGGATCAAT 61
Qy 442 AAATTCAGAGCTTGATCTCTTGAATAAAATCTTGAATTTGAAGGGCTTCATCGCCCT 501
Db 62 AAATTCAGAGATTGATACCTTATACAAATCTTGAATTTGAAGGGCTTCATAGGCCCT 121
Qy 502 CTTCTAGCTGTGCAGCTCACCAGCTCAAGGACGGGCTCACCATTGGATTAGCATTTAAC 561
Db 122 CTACTCGCGGTACAGCTCACCAGCTTAAAGAGCGGCTGGCAATGGATTAGCATTTAAT 181
Qy 562 CATGCTGCTCGATGGTACTTCGAGTGGGACTTTATGACCTCGTGGTCCAGCTTTCG 621
Db 182 CATGCTGCTCGAGCGGTACTTCAGCTGGGACTTCATGACTTCTCGGCCAGCTTTGT 241
Qy 622 TGTGGGTCCACCTCAATTTCTGTCACCATCTTGAACGAACCAAGGCTCGTAACACT 681
Db 242 AGTGGGCCACCTCCATCTCGTCCGCTCCACCTTCTCGAAAGAACTTAAGCTCGTACACT 301
Qy 682 CGAGTCAAGCTCAACCTCTCTCAACCATCAGATGCAACCGAATGCTTAAGTCAGCAACC 741
Db 302 CGAGTCAAACTCAACCTATCTAAGCCATCGGACGCGCGGACATGCCAAGTCGGAACG 361
Qy 742 AAGGTGATGTCGCGGCAAGCTAGACCCACCTCTTCGCGAAGAGATTTCAAGTTCTCC 801
Db 362 AAGCGGCGGTAGCGCCAGCGGTGATCCACCTATGCGTACAGGGGTGTTCAAGTTTTC 421
Qy 802 GAGTTAGCAATGACAAATCAAGTCAACAGTCAATGCCAATCAGGAGA----- 851
Db 422 GAATCAGCAATGATCAATCAAGTCAAAAGTCAATTAACCCAGGAGAGCGGCAAT 481
Qy 852 -----GACGCCATTTCCACATTCACATCTCCGCGACAGTGTGGCTAGCGCTCACA 906
Db 482 AATACGACGCGTCTCTACTTTCCATCCCTCTCTCCGCGACAGTGTGGCTAGCTGTACG 541
Qy 907 CGTCCGCGCAACTCAAGCGGAGGACTACACTGTGTACACTGTGTTGCTGATTCAGG 966
Db 542 CGTCCGAGGATCTCAAAACGCGAGGAATACAGGTTTACACCGGTTGTCGCTGATTCGCGT 601
Qy 967 AAAAGGTTGATCTCCATGTCAGAAAGTACTTCGCGCAACCTAAT 1014
Db 602 AAAGAGTCTGATCCGCGATGCGGAGAGTTACTTCGGAACCTAAT 649

RESULT 4
LOCUS BF051093 642 bp mRNA EST 16-OCT-2000
DEFINITION EST436268 tomato developing/immature green fruit Lycopersicon
esculentum cDNA clone cLEM21F13 5' sequence, mRNA sequence.
ACCESSION BF051093
VERSION BF051093.1 GI:10804989
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 642)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, immature green
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu.
Location/Qualifiers
1. .642

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEM21F13"
/clone_lib="tomato developing/Immature green fruit"
/tissue_type="fruit"
/dev_stage="immature green (5-35 days post-anthesis)"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCuadapt; Site_1: EcoRI;
Site_2: XhoI; Fruit were tagged at 5 dpa (0.5 cm) and
harvested at 7 day intervals through 35 dpa. Equal masses
of tissue from each stage were combined (including seeds
and locules) prior to mRNA isolation."
BASE COUNT 179 a 128 c 197 g 138 t
ORIGIN
Query Match 25.2%; Score 404.2; DB 143; Length 642;
Best Local Similarity 78.5%; Pred. No. 2.7e-89;
Matches 503; Conservative 0; Mismatches 123; Indels 15; Gaps 1;
Qy 715 GCACCGAATGCTAAGTTCAGCAACCAACGCTGATGTCGCCGCCAACGTAGACCCACCT 774
Db 2 GCGCCGAATGCTCCAAAGTCGGAACGAACGCGACGTTAGCGCCAGCGGTGATCCACT 61
Qy 775 CTTGCGCAAGAGTATTCAGATTCTCCGAGTTAGCAATTTGACAAAATCAAGTCAACAGTC 834
Db 62 ATGCGTGACAGGGTGTCAAGTTTCCGAATCAGCAATTTGATCAATCAAGTCAAAAGTC 121
Qy 835 AATGCCAATCAGGAGA-----GACGCCATTTCCACATTTCCCAATCCTC 879
Db 122 AATACTAACCCAGGAGAGAGCGCAATAATACGACGCGGTCTCTACTTTCCAATCCCTC 181
Qy 880 TCCGCACACGTGTGGCTAGCGCTCACACGTGCGCGCAACTCAAGCCCGAGGACTACACT 939
Db 182 TCGCCACAGTGTGGCTAGCTGTACGCGTGGCAGGAGCTCAACCGGAGGAATACAG 241
Qy 940 GTGTACACTGTGTGCTGCTGATTCGAGGAAAAGGTTGTCTCTCAATGCCAGAAAGTTAC 999
Db 242 GTTACACCGTGTGCTGCTGATTTGCCGTAAAGAGTGTGATCCGCCGATGCCGAGAGTTAC 301
Qy 1000 TTGCGCAACCTAATTCAGCAATTTTCACAGTGACCGCGCAGGTTTGTGTACTAGCAAGC 1059
Db 302 TTGGGAAACCTAATTCAGTCGATTTTCACTGTAAACGCGCGGGGATTTGCTGTATCGAAT 361
Qy 1060 CCGATCGAGTTCGCTGCTGGATGATACAAACAGCCATCGTGAAGCATGACGCTAAGGCC 1119
Db 362 CCGATTGAGTTTCCGCGACGGATGATTCATCAAGCCATAGCGAAGCAGCATGCGAAGCGG 421
Qy 1120 ATTGATGAAAGAAACAGGAGTGGGAGAGCAACCCGAAGATCTTTTCAGTACAAAGATGCT 1179
Db 422 ATTGAAGAAGGAACAAAGGAATGGGAGAGTAGTCCGAAGATATTTGCGTACAAAGATGCC 481
Qy 1180 GGAGTGAACCTGTGCTGCTGTTGGAAGTTTCGCCAAGGTTTCAAGGTTTACGACGTTT 1239
Db 482 AGAGTCAACTGTGTGCGGTGGGAGCTTCGCCAAGGTTTAAAGGAGTATGAGGTTGATTTT 541
Qy 1240 GGATGGGAAAGCCAGAGACTGTGAGAGTGTTCGAACAANTAGGTTTGTATGAANTGGTG 1299
Db 542 GGATGGGAAAGCCGAGATTCGCGAGGAGTGGGTTGAACAATAGATTTGACGCGGATGGTG 601
Qy 1300 TATTTCTACCAAGGCAAAATGAGGAGGAAGCAATGATGATGT 1340
Db 602 TATTTGTATCCAGAGCAAAATGGAGGAAGGAATTTGATGT 642
RESULT 5
LOCUS BE919634 588 bp mRNA EST 02-OCT-2000
DEFINITION BE919634 potato leaves and petioles Solanum tuberosum cDNA clone
cSTB2C8 5' sequence, mRNA sequence.
ACCESSION BE919634
VERSION BE919634.1 GI:10445710


```

182 ATCAACTTCTCGAAATTTGGAAAAAGATGATGAAGGAGTATTTAAAGGTGGAATATGATG 241
QY 359 ATGACATGGATGGTGTAGAGGTGCACAGTGGCTGTTCGACAGAGATAGAAAGTTGCAGATC 418
Db 242 ATGATATGGATGGATGTAGGTTAAATTTGTGGATGAGGCTCAAGAAATTCAGTTGCTTAATC 301
QY 419 TTAAGTCTG---ATGAAGAAGCACCACCAATTCGAGGACTTGTATTCCTGTCTTAATAAATCT 475
Db 302 TTAAGTCTGATCATGAAGGATCAATAAATTCGAAGATTTGATACCTTTATAACAAATCT 361
QY 476 TGAATTTGAAGGGCTTCATCGCCCTCTTCTAGCTGTGCAGCTCAACCAAGCTCAAGGACG 535
Db 362 TGAATTTGAAGGGCTTCATAGGCTCTACTCGCCGTACAGCTCAACCAAGCTTAAAGACG 421
QY 536 GGCTCACCATGGATAGCATTTAACCATGCTGTGTGCTGATGGTACTTCGACGTGGCACT 595
Db 422 GGCTGCAATGGATAGCATTTAATCATGCTGTGTGACGGTACTTCCACGTGGCACT 481
QY 596 TTATGACCTCGTGGTCCGAGCTTTGCTGTGGGTCCACCTCAATTTCTGTCCACCAATTC 655
Db 482 TCATGACTTCTGGGCCAGCTTTGTAGTGGGCCACCTCCATCTCGTCCGCTCCACCTTCC 541
QY 656 TTGAAGCAACCAAGGCTGTAAACATCGAGTCGAGTCAAGCTCAACCTCTCTCAACCATCAGATG 715
Db 542 TCGAAGAAGCACTAAGCTCGTGACATCGAGTCGAGTCAAACTCAACCTATCTAAGCCATCGGACG 601
QY 716 CACCCCAACATGC 728
Db 602 CGCCCGGAACATGC 614

RESULT 7
BG129408
LOCUS
DEFINITION
EST475054 tomato shoot/meristem Lycopersicon esculentum cDNA clone
CTOF24G12 5' sequence, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 651)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Utterback,T.,
Hansen,C., Ronning,C. and Tanksley,S.
Generation of ESTs from tomato shoot/meristem tissue
Unpublished (2001)
Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: http://www.genome.clemson.edu/orders/index.html.
Location/Qualifiers
1..651
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CTOF24G12"
/clone.lib="tomato shoot/meristem"
/tissue.type="shoot/meristem"
/dev_stage="developing shoots from 4-6wks old plants"
/lab_host="SOUR"
/note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
XhoI; Small expanding leaves from the growing tip were
taken from greenhouse plants (4-6wks old TA496). Tissue
was immediately frozen in liquid nitrogen."
203 a 123 c 151 g 174 t

BASE COUNT
ORIGIN
```

```

Query Match 24.1%; Score 386.8; DB 174; Length 651;
Best Local Similarity 79.8%; Pred. No. 5.5e-85;
Matches 519; Conservative 0; Mismatches 122; Indels 9; Gaps 5;

QY 122 CAAACCTGAAACACATATAACCACTAGGAAAAAAG-AGTGTCAATTTGGTAACATTTGAT 180
Db 1 CAAATCTGAAGCCACAAAACCACTAGGGAAAAAAGAAATGTCAATTAGTAACATTTGAT 60
QY 181 CTTCCCTTACCTAGCTTTCTATTACAACCAAAAATTTTCTCATCTATAA---AGTGTCTGAA 237
Db 61 TTACCTTTACTTACCTTTTATTATNACCAAAAGTTGATGTTTACAAAGTTAGGGGTGAA 120
QY 238 AACTTTGACGAGACGCTGGAAAAAATTAAGATGGACTGGCCCTTAGTTAGTTGGTGGATTC 297
Db 121 AGCTTTGACGAACACATAGAAAAATTTGAAGATGGTTTGGCTTTTAGTTGGAAGATTTT 180
QY 298 TATCAACTAGCTGGGAAACTTTGGAAAGATGAAGAAGGGTTTTTCAGGGTGGAAATACGAC 357
Db 181 TATCAACTTGTCTGAAAAATTTGGAAAAAGATGATGAAGGAGTATTTAAGGTGGAAATATGAT 240
QY 358 GATGACATGGATGGTGTAGAGGTGACAGTGGCTGTTCGAGAAGAGATAGAAAGTTGCAGAT 417
Db 241 GATGATATGGATGGAGTTTGAGGTTAATTTGTGATGAGGCTCAAGAAATTTCAAGTTGCTAAT 300
QY 418 CTTACTG---ATGAAGAAG3CACCAACCAATTCGAGGACTTGTATTCCTTGTGTAATAAATC 474
Db 301 CTTACTGATCATCATGAAGSGATCAATAATTCGAAGATTTTGATACCTTATAACAAATC 360
QY 475 TTGAATTTGGAAGGGCTTCATCGCCCTCTTCTAGCTGTGCAGCTCACCAGCTCAAGGAC 534
Db 361 TTGAATTTGAAGGGCTTCATAGGCTCTACTCGCCGTACAGCTCACCAGCTTAAAGAC 420
QY 535 GGCTCACCATGGGATTAGCATTTTAACCATGCTGTGCTGGATGGTACTTCGACGTGGCAC 594
Db 421 GGCTCGGCAATGGGATTAGCATTTTAATCATGCTGTGCTGGACGGTACTTCCACGTGGCAC 480
QY 595 TTTATGACCTCGTGGTCCGAGCTTTTCTGTGGTCCACCTCAATTTCTGTGCCACCATTC 654
Db 481 TTCATGACTTCTGGGCCACGCTTTGTAGTGGGCCACCTCCATCTCGTCCGACCTTTC 540
QY 655 CTTGAACGAACCAAGCTCTGTAACACTCGAGTCAAGCTCAACCTCTCTCAACCATCAGAT 714
Db 541 CTCGAAG-ACTAAAGCTCGTGACACTCGAGTCAAACTCAACCTATCTAAGCCATCCGAC 599
QY 715 GCACCCGACATGCTAAGTCAGCAACCAACGCGTGATGTCGCCGCCAACGT 764
Db 600 GCGCCGGAACATGCCAAGTC-GGAACAACGGCGACGTTAGCGCCACGCT 648

RESULT 8
AW618433
LOCUS
DEFINITION
EST320419 L. pennellii trichome, Cornell University Lycopersicon
pennellii cDNA clone cLPT13E7 5', mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Lycopersicon pennellii.
Lycopersicon pennellii
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 582)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Lakey,J., Holt,I.E.,
Liang,F., Hansen,T.S., Upton,J., Ronning,C.M., Craven,M.B., Fujii
,C.Y., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin
,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from wild tomato (Lycopersicon pennellii)
trichomes
Unpublished (1999)
Contact: David Frisch
```

Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.

FEATURES

source
Location/Qualifiers
1..582
/organism="Lycopersicon pennellii"
/db_xref="taxon:28526"
/clone="cuPTJ3E7"
/clone_lib="L. pennellii trichome, Cornell University"
/tissue_type="trichome"
/dev_stage="mixed stages"
/lab_host="SOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; Leaves of various stages were shaken in liquid
nitrogen, shearing off trichomes. This procedure yielded a
mixture of cells highly enriched for trichomes, with minor
contamination by other types of leaf cells."

BASE COUNT 176 a 109 c 135 g 162 t
ORIGIN
Query Match 23.1%; Score 371.4; DB 118; Length 582;
Best Local Similarity 79.9%; Pred. No. 3.4e-81;
Matches 464; Conservative 0; Mismatches 111; Indels 6; Gaps 2;
Qy 156 AGAGTGTCAATTTGGTAACATTTGATCTTCCTTACCTAGCTTTCTATTATACACCAAAATTT 215
Db 2 AGAATGTCAATTTAGTAACTTTGATTTACCTTACTTAGCTTTTATTATATACCAAAAGTT 61
Qy 216 TCTCATCTATAA---AGCTGCTGAAACCTTTGACGAGACGGTGGAAATAATTAAGATGG 272
Db 62 GATGGTTTACAAATTTAGGGCTGAAAGCTTTGAGGAAACAGTAGAATAATTTGAAAGATGG 121
Qy 273 ACTGGCCCTTAGTTGGTGGATTTCTATCAACTAGCTGGGAACTTGGAAAGATGAAGA 332
Db 122 TTGGCTTTAGTTGGAGATTTTATCAACTGCTGGGAAATTTGGAAAGATGATGA 181
Qy 333 AGGGTTTTTCAGGTGGGAATACGACGATGACATGGATGGTGTAGAGTGCACAGTGGCTGT 392
Db 182 AGGAGTATTTAAGGTGGAATATGATGATGATGATGATGATGATGATGATGATGATGATG 241
Qy 393 TGCAGACAGATAGAGTTGACAGATCTTAC---TGATGAAGAGCCACCACCAATTCGA 449
Db 242 GGCTCAAGAGATTCAAGTTGCTTAATCTTACTTAATCATCAAGAGGGGATTAAGAAATTCGA 301
Qy 450 GGACTTGATTCCTTGTATAAAATCTTGAATTTGGAAGGGCTTCATCGCCCTCTTCTAGC 509
Db 302 AGATTTGATACCTTATACAAAATCTTGAATTTAGAAGACTTCACAGGCTCTACTCGC 361
Qy 510 TGTGAGCTCACCAGCTCAAGGACGGGCTCACCATGGGATTTAGCATTTAACCATGCTGT 569
Db 362 CGTACAGCTCACCAGCTTAAAGAGGGGCTGGCAATGGGATTTAGCATTTAATCATGCTGT 421
Qy 570 GCTGGATGGTACTTCGAGTGGCAGCTTTATACCTCTGGTCCGAGCTTTCCTGTGGGTC 629
Db 422 GCTGGAGGGTACCTCGACGTGGCAGCTTTCATGACTTCTTGGGCCAGCTCTCTAGTGGGGC 481
Qy 630 CACCTCAATTTCTGTCCACCACCTTCTTGAAGCAACCAAGGCTCGTAACTTCGAGTCAA 689
Db 482 CACCTCCATCTCGGTCCACCTTCTTCCGACGAGCACTTAAGGCTCTGTACACTCGAGTCAA 541
Qy 690 GCTCAACCTCTCTCAACCATCAGATGCACCCGGAACATGCTA 730
Db 542 ACTCAACCTATCTAAGCCATCGCAGCGACCGGAACATGCCA 582

RESULT 9

BE920362

LOCUS

, BE920362 569 bp mRNA EST 02-OCT-2000

DEFINITION

EST424131 potato leaves and petioles Solanum tuberosum cDNA clone

CSTB5P22 5' sequence, mRNA sequence.

ACCESSION BE920362

VERSION BE920362.1 GI:10446438

KEYWORDS EST.

SOURCE potato.

ORGANISM

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

1 (bases 1 to 569)

van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,

Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning

,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.

Generation of ESTs from potato leaves and petioles

Unpublished (2000)

TITLE

Contact: Cathy Ronning

The Institute for Genomic Research

For clone request: please contact Research Genetics, Libraries

Division tel 1-800-711-6195, email cdna@resgen.com.

FEATURES

Location/Qualifiers

1..569

/organism="Solanum tuberosum"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="CSTB5P22"

/clone_lib="potato leaves and petioles"

/tissue_type="leaflets and petioles"

/dev_stage="8 weeks old plants"

/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

XhoI; Tissue was supplied by Dr. Fry (Cornell University).

Leaflets and petioles were isolated from 8 week old

greenhouse grown plants. The plants were watered and

fertilized freely. The tissue was immediately frozen in

liquid nitrogen."

BASE COUNT 186 a 84 c 136 g 163 t

ORIGIN

Query Match 22.7%; Score 363.6; DB 142; Length 569;
Best Local Similarity 82.3%; Pred. No. 2.9e-79;
Matches 443; Conservative 0; Mismatches 89; Indels 6; Gaps 2;
Qy 96 TACAAAAGTGAAGTCTCGAAAAAACAACAGCTGAACACCATAAACACCATAGGAAAAA 155
Db 32 TACAAATGTGAACATCTTTGAAAAAATCAATGTGAAGCTCAAAACACCATAGGAAAAA 91
Qy 156 AGAGTGTCAATTTGGTAAACATTTGATCTTCCTTACCTAGCTTTCTATTACACCAAAAAAT 215
Db 92 AGAATGTCAATTTGGTAAACATTTGATTTACCTTACTTACTTTTATTATACCAAAAGTT 151
Qy 216 TCTCATCTATAA---AGGTGCTGAAACCTTTGACGAGACGGTGGAAAAAATTAAGATGG 272
Db 152 GATGGTTTACAAATTTAGGGACTGAAAACCTTTGAGAAACAGTGGAAAAATTTGAAAGATGG 211
Qy 273 ACTGGCCTTACTATTGGTGGATTTCTATCACTAGCTGGAAACTTGGAAAAGATGAAGA 332
Db 212 TTGGCTTTAGTTGGGAAGATTTTATCACTGCTGGAAATTTAGGAAAAGATGAAGA 271
Qy 333 AGGGTTTTTCAGGTGGGAATACGACGATGACATGGTGTAGAGGTGACAGTGGCTGT 392
Db 272 AGGAGTATTTAAGTGGGAATATGATGATATGATGATGATGATGATGATGATGATGATG 331
Qy 393 TGCAGAGAGATAGAAGTTGCAGATCTTACTG---ATGAAGAAGGCCACCAACCAATTCGA 449
Db 332 AGCTGAAGGGATTGAAGTTGCAAGTCTTACTGATCATCAAGAAGGGATTAATAATTCGA 391
Qy 450 GGACTTGATTCCTTGTATAAAATCTTGAATTTGGAAGGGCTTCATCGCCCTCTTCTAGC 509
Db 392 AGCTTTGATACCTTATACAAAATCTTGAATTTGGAAGGGCTTCATAGGCTCTTCTCGC 451
Qy 510 TGTGACGCTCACCAAGCTCAAGGACGGGCTCACCATGGGATTAGCATTTAACCATGCTGT 569

Db	452	CGTTGAGCTCACCAGCTCAAAAGACGGCTGGCAATGGGATAGCATTTAATCATGCTGT	511
Qy	570	GCTGGATGGTACTTCTGACGTGGCAGCTTTATGACCTCGTGGTCCGAGCTTTGCTGTGGG	627
Db	512	GCTGGATGGTACTTCCACGTCGCACCTTCATGACTTCATGGCCCGCCAGCTTTGATGGG	569
RESULT	10		
AW039511			
LOCUS		561 bp mRNA EST	18-OCT-1999
DEFINITION		EST281792 tomato mixed elicitor, BTI Lycopersicon esculentum cDNA clone cLET14C2, mRNA sequence.	
ACCESSION		AW039511	
VERSION		AW039511.1	GI:5898265
KEYWORDS		EST.	
SOURCE		tomato.	
ORGANISM		Lycopersicon esculentum	
REFERENCE		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;	
AUTHORS		Lycopersicon. 1 (bases 1 to 561) D'Ascenzo,M., He,X., Lyman,J., Holt,I.E., Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.F., Bowman,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,S.D. and Giovannoni,J.	
TITLE		Generation of ESTs from tomato leaf tissue	
JOURNAL		Unpublished (1999)	
COMMENT		Contact: David Frisch Clemson University Genomics Institute 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrische@LEMON.EDU 5 prime sequence.	
FEATURES		Location/Qualifiers	
source		1..561	
		/organism="Lycopersicon esculentum"	
		/cultivar="Rio Grande Ptor"	
		/db_xref="taxon:4081"	
		/clone_lib="tomato mixed elicitor, BTI"	
		/clone="cLET14C2"	
		/tissue_type="leaf"	
		/dev_stage="4-6 week old plants"	
		/lab_host="XL1-Blue MRP"	
		/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLET - inoculated with a variety of disease response elicitors. Plants exposed to 2,6 dichloroisonicotinic acid, BTH, jasmonic acid, ethylene, fenthion, EIX, okadaic acid, or systemin prior to tissue harvest. EcoRI site was destroyed during cloning."	
BASE COUNT		184 a 87 c 129 g 161 t	
ORIGIN			
Query Match		22.18;	Score 355.4; DB 110; Length 561;
Best Local Similarity		80.58;	Pred. No. 3e-77;
Matches	442;	Conservative	0; Mismatches 101; Indels 6; Gaps 2;
Qy	96	TACAAAGTGAAGTCTCTGAAAAAACAAAGCTGTAACACCATATAACCACTAGGAAAAA	155
Db	13	TACAAATGTGAACATCTTGAAAAAATCAATGTAAGGCCAACAAACCACTAGGAAAAA	72
Qy	156	AGAGTGTCAATTGGTAACATTTGATCTTCCTTACCTAGCTTTCTATTACAACCAAAAT	215
Db	73	AGAATGTCAATTAGTAACATTTGATTTACCTTACTTAGCTTTTATATAACCAAAAGT	132
Qy	216	TCTCATCTATAA---AGTGTCTGAAACATTTTCACGAGACGGTGGAAAAAATTAAGATGG	272
Db	133	GATGGTTTACAAGTTAGCGGCTGAAAGCTTTTGAGGAAACAGTAGAAAAAATTTGAAGATGG	192

ORIGIN

Query Match 22.1%; Score 355.2; DB 119; Length 544;
Best Local Similarity 80.1%; Pred. No. 3.4e-77;
Matches 436; Conservative 0; Mismatches 93; Indels 15; Gaps 1;

QY 487 GGGCTTCATCGCCCTCTTCTAGCTGTGAGCTCACCAGCTCAAGGAGCGGCTCACCATG 546
Db 1 GGGCTTCATCGCCCTCTTCTAGCTGTGAGCTCACCAGCTCAAGGAGCGGCTCACCATG 60

QY 547 GATTAGCATTAACCATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 606
Db 61 GATTAGCATTAACCATGCTGTGCTGATGCTGATGCTGATGCTGATGCTGATGCTGATGCTG 120

QY 607 TGGTCCGAGCTTGTGCTGGGTCACCAATTTCTGCTCCACCAATTTCTGCTCCGCAAC 666
Db 121 TGGGCCAGCTTGTAGTGGGCGCCACCTCCATCTCGGTCCCACTTCTCTGAAAGAACT 180

QY 667 AAGGCTCGTAACATCGAGTCAAGCTCAACCTCTCTCAACCATCAAGTCAAGTCAAGTCA 726
Db 181 AAGGCTCGTAACATCGAGTCAACCTCAACCTCTCTCAACCATCAAGTCAAGTCAAGTCA 240

QY 727 GCTAGTCAGCAACCAAGGCTGATGCTCGGCGCAACGCTAGACCCACCTCTCTCGGCAAGA 786
Db 241 GCAAGTCGGAACCAAGGCGAGCTTAGCGCGCAGCTGATGCTGATGCTGATGCTGATGCTG 300

QY 787 GTATTCAAGTTCTCCGAGTACCAATTCACAAATTCAGTCAACGATCAACGATCAACGATCA 846
Db 301 GTATTCAAGTTCTCCGAGTACCAATTCAGTCAACGATCAACGATCAACGATCAACGATCA 360

QY 847 GGAGA-----GACGCGATTTCTCCACATTTCCACATTTCCGCGCACACGCTG 891
Db 361 GGAGAAGGCGCAATTAATACAGCGCTTCTCTACTTTTCCATCTCTCTCCGCGCACGCTG 420

QY 892 TGGCTAGCGTCACAGCTGCGGCGCACTCAAGCCGAGGACTACACTGTGTACACTGTG 951
Db 421 TGGCTAGCTGTACAGCTGCGGCGCAGCTCAACCGGAGGATACACGCTTTTACACCGCTG 480

QY 952 TTTGCTGATTGCGAGAAAGGTTGATCTTCAATTCGAGAAAGTTACTTCCGCAACCTA 1011
Db 481 TTCGCTGATTGCGGTAAGAGAGTCATCCGCGGATGGCGGAGAGTTACTTCCGAAACCTA 540

QY 1012 ATTC 1015
Db 541 ATTC 544

RESULT 12
BF053244
LOCUS BF053244 556 bp mRNA EST 16-OCT-2000
DEFINITION EST438474 potato leaves and petioles Solanum tuberosum cDNA clone
CSTB34L5 5' sequence, mRNA sequence.
ACCESSION BF053244
VERSION BF053244.1 GI:10807140
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 556)
van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Llang, F., Cho, J.,
C.M., Fry, W.E., Tanksley, S.D. and Baker, B.
Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
Location/Qualifiers
FEATURES
* source 1..556

/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTB34L5"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 185 a 83 c 130 g 158 t
ORIGIN

Query Match 22.0%; Score 353.8; DB 143; Length 556;
Best Local Similarity 82.3%; Pred. No. 7.6e-77;
Matches 432; Conservative 0; Mismatches 87; Indels 6; Gaps 2;

QY 96 TACAAAGTGAAGTCTCGAAAAAACAACGCTGAACCAACATTAACCACTAGGAAAAA 155
Db 32 TACAAATGTGAACATCTTGAATAATCAATGTGAAGCTCAAAAACCACTAGGAAAAA 91

QY 156 AGAGTGTCAATTTGGTAACATTTGATCTTCTTACCTAGCTTTCTATTACACCAAAAT 215
Db 92 AGATGTCAATTTGGTAACATTTGATCTTCTTACCTAGCTTTCTATTATTAACCAAAAT 151

QY 216 TCTCATCTATAA---AGGTGCTGAAACCTTTGACGAGACGGTGGAAAAAATAAAGATGG 272
Db 152 GATGTTTACAAATTTAGGAGCTGAAACCTTTGAGGAAACAGTGGAAAAATTTGAAAGATGG 211

QY 273 ACTGCCCTTAGTATTGGTGGATTTCTATCACTAGCTGGGAACTTGGAAAGATGAAGA 332
Db 212 TTTGGCTTTAGTGTGGGAAGATTTTATCACTTCTCGGAAATTTAGGAAAAAGATGAAGA 271

QY 333 AGGGTTTTCAGGGTGGAAATACGAGATGACATGCTGTAGAGGTGACAGTGGCTGT 392
Db 272 AGGATTTTAAAGTGGAAATATGATGATGATGATGATGATGATGATGATGATGATGATG 331

QY 393 TGCAGAAGAGATAGAAGTTGCAGATCTTACTG---ATGAAGAAGGCACCAACAAATGCA 449
Db 332 AGCTGAAGGGATTTGAAGTTGCAAGTCTTACTGATCATCAAGAAGGGATTAATAAATCCA 391

QY 450 GGACTTGATCTCTGTATAAATTTGCAATTTGGAAGGGCTTCATCGCCCTCTCTCTAGC 509
Db 392 AGCTTTGATACCTTATAAATAATCTTGAATTTGGAAGGGCTTCATAGGGCTCTCTCTCGC 451

QY 510 TGTGAGCTCACCAGCTCAAGGAGCGGCTCACCATGGGATTAGCATTTAACCATGCTGT 569
Db 452 GTTCAAGCTCACCAGCTCAAGGAGCGGCTGCAATGGGATGAGCATTTAATCATGCTGT 511

QY 570 GCTGATGGTACTTTCAGCTGGCAGCTTTATGACCTTCGCTGGTCCGA 614
Db 512 GCTGATGGTACTTTCAGCTGGCAGCTTTATGACCTTCATGACCTTCATGCTGGGCCCA 556

RESULT 13
BF275133
LOCUS BF275133 786 bp mRNA EST 07-MAR-2001
DEFINITION GA_EB0023D13f Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_EB0023D13f, mRNA sequence.
ACCESSION BF275133
VERSION BF275133.1 GI:11206203
KEYWORDS EST.
SOURCE Gossypium arboreum.
ORGANISM Gossypium arboreum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.
1 (bases 1 to 786)

nitrogen, shearing off trichomes. This procedure yielded a mixture of cells which is highly enriched for trichome cells" likely with minor contaminations of other types of leaf cells"

BASE COUNT 133 a 153 c 135 g 124 t
ORIGIN

Query Match 20.7%; Score 332.2; DB 118; Length 545;
Best Local Similarity 77.4%; Pred. No. 1.6e-71;
Matches 422; Conservative 0; Mismatches 108; Indels 15; Gaps 1;

Qy 565 GCTGTGCTGGATGCTTTCGACGCTGGCAGCTTATATGACCTCTGCTGCTCGAGCTTTGCTGT 624
Db 1 GCTGTGCTGGACGCTTCTTCCACGCTGCACCTCATGACTTCTCTGGCCACGCTTTGTAGT 60

Qy 525 GGGTCCACCTCAATTTCTGTCCCAACCATTCCTTGAAGCAACCAAGGCTCGTAACACTCGA 684
Db 61 GGGGCCACCTCATCTCGGTCCCACTTCTCTCGACCGCACTAAAGCTCGTGACACTCGA 120

Qy 685 GTCAAGCTCAACCTCTCTCAACCATCAGATGCACCCGCAACATGCTAAGTCAGCAACCAAC 744
Db 121 GTCAACTCAACCTATCTAAGCCATCCGACCGCAACATGCCAAGTCGGAACCAAC 180

Qy 745 GGTGATGTCCCGGCCAACGTAGACCCACCTCTTCGCGAAAGAGATTTCAAGTTCTCCGAG 804
Db 181 GGCACATTAGCGCCAGCGTGGATCCACCTATGCTGACAGGGTGTTCAGTTTCCGAA 240

Qy 805 TTACCAATTTGACAAATCAAGTCAACAGTCAATGCCAACTCAGGAGA----- 851
Db 241 TCAGCGATTGATCAAAATCAAGTCAAAAGTCAATACGAATCCACCAGAGGCACAGATAAC 300

Qy 852 --GAGCGCATTTCTCCACATTCCTCAATCACTCTCCGACACAGTGTGCTAGCGTCACAGT 909
Db 301 AGTCTGCTGCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 360

Qy 910 GCGCGCAACTCAAGCCCGAGGACTACACTGTGTACACTGTGTGTGTGTGTGTGTGTGTGTGT 969
Db 361 GCGAGGCGAGCTGAAACCCGAGGAGTAATACAGGTTTACACCGTGTTCGCTGATTGCCGTA 420

Qy 970 AGGTTGATCTCTCAATGCCAGAAAGTTACTTTCGCAACCTTAATTCAGGCAATTTTCACA 1029
Db 421 AGAGTCGATCCGCGGATGCCGAGAGTTACTTTCGGAACCTTAATTCAGGCGATTTCCTACT 480

Qy 1030 GTGACCGCGGAGCTTTGTTTACTAGCAAGCCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1089
Db 481 GTACGCTCGGGGATTGCTGCTATCGAATCCGATTGAGTTTGGCGGAGGATGATTCAT 540

Qy 1090 CAAGC 1094
Db 541 GAAGC 545

Search completed: November 5, 2001, 18:00:51
Job time: 14891 sec

Db 31 TACAAATGTGAACATCTTGAATAAATCAAAATGTGAAGCCACAAAACCACTAGGGAATA 90

Qy 156 AGAGTGTCAATTTGGTAACATTTGCTTCCCTTACCTAGCTTCTTATTAACAACCAAAATTT 215
Db 91 AGAATGTCAATTTAGTAACATTTGATTTACCTTACTTACTTCTTATTAACCAAAATTT 150

Qy 216 TCTCATCTATAA---AGCTGCTGAAACTTTTACGAGACGCTGGAAATAAATAAAGATGG 272
Db 151 GATGCTTTTACAAGTTAGGGCTGAAAGCTTTGAGGAAACACGTAGAAAATTTGAAAGATGG 210

Qy 273 ACTGCCCTTAGTATTTGGGATTTCTATCAACTAGCTGGGAAACTTGGAAAAGATGAAGA 332
Db 211 TTGCTTTTAGTGTGGGAAGATTTTATCAACTTCTGGAATAATTTGGAAAAGATGATGA 270

Qy 333 AGGGTTTTTCAGGCTGGGAATACGACGATGACATGCTGTAGAGTGACAGTGGCTGT 392
Db 271 AGGAGTATTTAAGTGGGAATATGATGATATGATGGAGTTGAGGTAAATTTGGATGA 330

Qy 393 TGCAGAAGAGATAGAAGTTGCGAGATCTTACTGATGAA---GAAGGCCACCAAAATTTGCA 449
Db 331 GGCTCAAGAAATTCAGTTGCTAATCTTACTGATCATCATGAAGGATCAATAAATTTCCA 390

Qy 450 GGACTTGATCTCTGTATATAAATCTTGAATTTGGAAGGCTTATCGCCCTCTTCTAGC 509
Db 391 AGATTTGATACCTTATATAAATAATCTTGAATTTTGAAGGCTTCTATAGGCTCTACTCGC 450

Qy 510 TGTGAGCTCACCAAGCTCAAGGAGCGGCTCACCAGTGGGATTTAGCATTTAACCATGCTGT 569
Db 451 CGTACAGCTCACCAAGCTTAAAGAGGCTGGCAATGGGATAGCATTTAATCATGCTGT 510

Qy 570 GCTGATGCTACTTTCGAGTGGGCTTTATGACCTCGTGTGCTGCTGCTGCTGCTGCTGCTGCT 620
Db 511 GCTGAGCGTACTTCCAGTGGCACTTTCATGACTTCTCTGGGCCCGAGCTTTG 561

RESULT 15

AW617090 545 bp mRNA EST 24-MAR-2000
LOCUS EST323501 L. hirsutum trichome, Cornell University Lycopersicon
DEFINITION hirsutum cDNA clone cLHT21F17 5', mRNA sequence.

ACCESSION AW617090
VERSION AW617090.1 GI:7323200
KEYWORDS EST.
SOURCE Lycopersicon hirsutum.
ORGANISM Lycopersicon hirsutum
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.

REFERENCE 1 (bases 1 to 545)
AUTHORS van der Hoeven,R.S., Bezzarides,J.L., Matern,A.L., Holt,I.E., Liang
F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from wild tomato (Lycopersicon hirsutum)
trichomes
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
LOCATION/Qualifiers
1. .545
/organism="Lycopersicon hirsutum"
/db_xref="taxon:62890"
/clone="cLHT21F17"
/clone_lib="L. hirsutum trichome, Cornell University"
/tissue_type="trichome"
/dev_stage="mixed stages"
/note="Leaves of various stages were shaken in liquid

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:46:27 ; Search time 8904.87 Seconds
(without alignments)
2569.021 Million cell updates/sec

Title: US-08-894-356C-4
Perfect score: 1479
Sequence: 1 CCGTGATCGAACGCTGTAGA.....TTGAAAAAAAAAAAAAAAAAAAA 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

- Database : GenEmbl:*
- 1: gb_ba1:*
 - 2: gb_ba2:*
 - 3: gb_ba3:*
 - 4: gb_in1:*
 - 5: gb_in2:*
 - 6: gb_in3:*
 - 7: gb_om:*
 - 8: gb_ov:*
 - 9: gb_pat1:*
 - 10: gb_pat2:*
 - 11: gb_ph:*
 - 12: gb_pl1:*
 - 13: gb_pl2:*
 - 14: gb_pl3:*
 - 15: gb_pl4:*
 - 16: em_ba1:*
 - 17: em_ba2:*
 - 18: em_fun:*
 - 19: em_htgo_hum:*
 - 20: em_htgo_inv:*
 - 21: em_htgo_rod:*
 - 22: em_htg_hum1:*
 - 23: em_htg_hum2:*
 - 24: em_htg_hum3:*
 - 25: em_htg_hum4:*
 - 26: em_htg_hum5:*
 - 27: em_htg_hum6:*
 - 28: em_htg_hum7:*
 - 29: em_htg_hum8:*
 - 30: em_htg_inv1:*
 - 31: em_htg_inv2:*
 - 32: em_htg_other:*
 - 33: em_hum1:*
 - 34: em_hum2:*
 - 35: em_hum3:*
 - 36: em_hum4:*
 - 37: em_hum5:*
 - 38: em_hum6:*
 - 39: em_hum7:*
 - 40: em_in:*
 - 41: em_om:*
 - 42: em_or:*
 - 43:

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_rod:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vi:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_v1:*
- 59: gb_v2:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rod:*
- 95: gb_rod2:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1478.6	100.0	1479	10	E12756
2	1447.6	97.9	1476	12	AB029340
3	218.4	14.8	1622	10	E12754
4	218.4	14.8	1622	12	AB026494
5	190.6	12.9	1508	10	E12757
6	172.2	11.6	1679	12	AB010708
7	172.2	11.6	1703	10	E12753
8	110.8	7.5	66237	12	AB016892

- E12756 Perilla oci
- AB029340 Perilla f
- E12754 Gentianatri
- AB026494 Gentiana
- E12757 Senecio cru
- AB010708 Gentiana
- E12753 Gentianatri
- AB016892 Arabidops

9	110.8	7.5	101176	12	AC002560	Genomic s
10	99.6	6.7	119914	12	AC003027	Arabidops
11	76.8	5.2	1648	9	AX025514	Sequence
12	75.4	5.1	85690	12	AB028618	Arabidops
13	75.4	5.1	136047	74	AC069470	Arabidops
c 14	68.8	4.7	136047	74	AC069470	Arabidops
c 15	65.8	4.4	7218	10	I66494	Sequence 14
c 16	63.6	4.3	85690	12	AB028618	Arabidops
c 17	63.4	4.3	82360	13	AP000606	Arabidops
c 18	63.4	4.3	341064	75	AC074226	Arabidops
19	57.2	3.9	7218	10	I66494	Sequence 14
c 20	57	3.9	78379	12	AB006696	Arabidops
c 21	56.8	3.8	71623	77	AC087724	Arabidops
22	52.6	3.6	111338	13	ATF16L2	Arabidops
c 23	52	3.5	93127	74	AC069373	Homo sapi
c 24	52	3.5	186857	74	AC069336	Homo sapi
c 25	51	3.4	14553	4	AE001391	Plasmodiu
c 26	50.8	3.4	46739	13	AP002052	Arabidops
c 27	50.2	3.4	36075	6	CELT27A1	Caenorhab
c 28	49.4	3.3	253305	96	PFMAL3P7	Plasmodiu
29	49	3.3	167336	66	AC021350	Homo sapi
c 30	48.4	3.3	156060	60	AC004153	Plasmodiu
31	47.8	3.2	132781	62	AC011689	Homo sapi
32	47.8	3.2	144014	62	AC011685	Homo sapi
33	47.8	3.2	179060	75	AC074382	Homo sapi
c 34	47.2	3.2	42879	93	HSAC000123	Human Cos
c 35	47.2	3.2	45508	93	HSAC000110	Human Cos
c 36	47.2	3.2	245802	60	AC006279	Plasmodiu
c 37	47	3.2	759	53	CNS06QXV	T7 end of
c 38	47	3.2	12029	4	AE001430	Plasmodiu
c 39	47	3.2	63954	77	AC087789	Homo sapi
40	47	3.2	93196	67	AC022085	Homo sapi
41	47	3.2	169546	60	AC004157	Plasmodiu
c 42	47	3.2	180301	65	AC018362	Homo sapi
c 43	46.8	3.2	2437	5	AF006380	Tetrahyme
c 44	46.8	3.2	194264	90	AL356956	Human DNA
45	46.4	3.1	2452	5	AF337815	Dictyoste

ALIGNMENTS

RESULT	1
E12756	
LOCUS	Perilla ocimoides mRNA for acyltransferase, partial cds.
DEFINITION	Perilla ocimoides mRNA for acyltransferase, partial cds.
ACCESSION	E12756
VERSION	E12756.1 GI:3251588
KEYWORDS	JP 1997070290-A/4.
SOURCE	unidentified.
ORGANISM	unclassified.
REFERENCE	1 (bases 1 to 1479)
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T. .
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL	Patent: JP 1997070290-A 4 18-MAR-1997; SUNTOBY LTD
COMMENT	OS Perilla ocimoides PN JP 1997070290-A/4 PD 18-MAR-1997 PF 30-JAN-1996 JP 1996046534 PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHITO, PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAARI PC C12N15/09,A01H1/00,C07K14/42,C12N9/10,(C12N9/10, PC C12R1:865), PC (C12N9/10,C12R1:19); CC strandedness: Double; CC topology: Linear; FH Key Location/Qualifiers

FT	source	1. .11479	/organism='Perilla ocimoides' FT
FT	/tissue_type='leaves'		
FT	/clone='pSAT208'		
FT	CDS	1. .1343	
FT	/product='acyltransferase'.		
FEATURES	Location/Qualifiers		
source	1. .1479		
	/organism='unidentified'		
	/db_xref='taxon:32644'		
BASE COUNT	420 a	316 c	331 g 411 t 1 others
ORIGIN			
Query Match 100.0%; Score 1478.6; DB 10; Length 1479;			
Best Local Similarity 99.9%; Pred. No. 0;			
Matches 1478; Conservative 1; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	CCGTGATCGAAACGTGTAGAGTTGGCGCGCGCGGACTCGTGGCGGAGCAATCGGTGC	60
Db	1	CCGTGATCGAAACGTGTAGAGTTGGCGCGCGCGGACTCGTGGCGGAGCAATCGGTGC	60
Qy	61	CGTCACATTTCTGCACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCTCTTCT	120
Db	61	CGTCACATTTCTGCACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCTCTTCT	120
Qy	121	ACGAATTCCTTGTTCACAAACAATTTTTCAGAAATCCATCGTTCACAACTCAACAAT	180
Db	121	ACGAATTCCTTGTTCACAAACAATTTTTCAGAAATCCATCGTTCACAACTCAACAAT	180
Qy	181	CTCTCTCTAAACACTCTCATACACTTTCTCCCTCTCTCATGCAATTAATCTACCTTCAT	240
Db	181	CTCTCTCTAAACACTCTCATACACTTTCTCCCTCTCTCATGCAATTAATCTACCTTCAT	240
Qy	241	CCCGGAGAAAATCCGGAGTTTCGGTATCTATCCGGGAGCTCGGTTCTTTTCACCATCG	300
Db	241	CCCGGAGAAAATCCGGAGTTTCGGTATCTATCCGGGAGCTCGGTTCTTTTCACCATCG	300
Qy	301	CAGATCTAGCGAGACTTCGATGATCTCTCGGAAATCGCCAGAAATCCCGGTAGGC	360
Db	301	CAGATCTAGCGAGACTTCGATGATCTCTCGGAAATCGCCAGAAATCCCGGTAGGC	360
Qy	361	TCTACAACTTTTGCCTAAATTTGCCGCCATTGTGCGAAGAATCCGATAGAAACTCTTCC	420
Db	361	TCTACAACTTTTGCCTAAATTTGCCGCCATTGTGCGAAGAATCCGATAGAAACTCTTCC	420
Qy	421	AAGTTTTCGGCGTGCAGGTGACTCTTTTCCAGGCCGAGGCGTGGTATTGGAATAGCAA	480
Db	421	AAGTTTTCGGCGTGCAGGTGACTCTTTTCCAGGCCGAGGCGTGGTATTGGAATAGCAA	480
Qy	481	CGCATCACACCGTTAGCGACGCCGCTGTTCTCGCGTTTATACGGCTTGGTCTTCAA	540
Db	481	CGCATCACACCGTTAGCGACGCCGCTGTTCTCGCGTTTATACGGCTTGGTCTTCAA	540
Qy	541	TGAGCAACACATTGAAATGAAGATGAAGATGAAGATTTAAATCTTTTGGCAGTTTCG	600
Db	541	TGAGCAACACATTGAAATGAAGATGAAGATGAAGATTTAAATCTTTTGGCAGTTTCG	600
Qy	601	ATAGATCGGTCAATAATATCCGAGAAATTTGACTTCATTTATTTGGAGAAACGGCTAA	660
Db	601	ATAGATCGGTCAATAATATCCGAGAAATTTGACTTCATTTATTTGGAGAAACGGCTAA	660
Qy	661	AATTTCTTTCGATCTCGTATCCCTCATATCCGAGCCGCGGATTCGACACCGGTTTCG	720
Db	661	AATTTCTTTCGATCTCGTATCCCTCATATCCGAGCCGCGGATTCGACACCGGTTTCG	720
Qy	721	TTTTTACCCAAATCCAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT	780
Db	721	TTTTTACCCAAATCCAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAATTAAGAAAT	780
Qy	781	TAGTCCATCTCTCTCTTTTGTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAGG	840
Db	781	TAGTCCATCTCTCTCTTTTGTAGGATTTAGGATTTAGGATTTAGGATTTAGGATTTAGG	840

FEATURES		Location/Qualifiers	
Source		1..1476	
		/organism="Perilla frutescens"	
		/db_xref="taxon:48386"	
		/tissue_type="leaf"	
		<1..1:43	
CDS			
		/codon_start=3	
		/product="anthocyanin acyltransferase"	
		/protein_id="BA93475.1"	
		/translation="VIETCRVGGPPDPSVAEQSVPLTFDFDWTMLHFHPMLQLLFYEPFC	
		SQHFSSEIVGNPKLQKSLKTLHFFPLSLNLIYPPSPKPMPEFRYLSGDSVSPTIAES	
		SKDFDLVGNRPESPVRLNFVKLPPIVEESDRKLFQVFAVOTLFPGRGVGIGIAT	
		HITVSDAPSFLAFITAWSSMSKHIEDEDEPKSLPVFDRSVIKYPTKFDSTLYWRNA	
		LKEPLQSRHPSLPTDRIRTFVFTQSKIKKLGWIOQSRVPSLVHLSSFVAIAAYMWAG	
		ITKSFADQDQDNEDAFFLIPVDLRPLDPPVPENFTGNCLSLSTALPRMRRELVLGKGM	
		VFAAEVIAAEIKKRINDLETVEKSPKEIRKALOKYSFVAGSKLDLVLGADFGW	
		GKARKQELSIDGEKYAMTLCKARDFEGGLEVCLSLPKDKMDAFAYFSLGNG"	
BASE COUNT		419 a 315 c 331 g 411 t	
ORIGIN			
Query Match		97.9%; Score 1447.6; DB 12; Length 1476;	
Best Local Similarity		99.6%; Pred. No. 0;	
Matches 1472; Conservative		0; Mismatches 4; Indels 2; Gaps 2;	
QY	1	CCGTGATCGAAAGCTGTAGAGTTGGGCCGCCGCCGACTCGGTGGCGGAGCAATCGGTGC	60
DB	1	CCGTGATCGAAAGCTGTAGAGTTGGGCCGCCGCCGACTCGGTGGCGGAGCAATCGGTGC	60
QY	61	CGCTCACATTTCTTCGACATGACGTGGCTGCTATTTTCATCCCATGCTTCAGCTCCTCTTCT	120
DB	61	CGCTCACATTTCTTCGACATGACGTGGCTGCTATTTTCATCCCATGCTTCAGCTCCTCTTCT	120
QY	121	ACGAATTCCTTGTTCGACGCAACATTTTTCAGAAATCCATCGTTCCAAAACCTCAACAAT	180
DB	121	ACGAATTCCTTGTTCGACGCAACATTTTTCAGAAATCCATCGTTCCAAAACCTCAACAAT	180
QY	181	CTCTCTTAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTAATCTACCCCTTCAT	240
DB	181	CTCTCTTAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTAATCTACCCCTTCAT	240
QY	241	CCCGGAGAAATCGCGGAGTTTCGGGTATCTATCCGGGGACTCGGTTCCTTTTCCACCATCG	300
DB	241	CCCGGAGAAATCGCGGAGTTTCGGGTATCTATCCGGGGACTCGGTTCCTTTTCCACCATCG	300
QY	301	CAGAAATCTAGCGAGCTTCGATGATCTCGTGGAAATCGTCGAAATCTCCCGTTTAGCG	360
DB	301	CAGAAATCTAGCGAGCTTCGATGATCTCGTGGAAATCGTCGAAATCTCCCGTTTAGCG	360
QY	361	TCACAACTTTGTCCCTAAATTCGCCGCCCATTTGCGAAGAAATCCGATAGAAAACCTCTCC	420
DB	361	TCACAACTTTGTCCCTAAATTCGCCGCCCATTTGCGAAGAAATCCGATAGAAAACCTCTCC	420
QY	421	AAGTTTTCGCGCGTCGAGGTGACTTCTTTTCCAGGCCGAGGCGTGGGTATTTGGAATAGCAA	480
DB	421	AAGTTTTCGCGCGTCGAGGTGACTTCTTTTCCAGGCCGAGGCGTGGGTATTTGGAATAGCAA	480
QY	481	CGCATACACCGTTAGGAGCGCCCGTGGTTCTCGGGTTTATAACGGCTTGGTCTTCAA	540
DB	481	CGCATACACCGTTAGGAGCGCCCGTGGTTCTCGGGTTTATAACGGCTTGGTCTTCAA	540
QY	541	TGAGCAACACATTGAAATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT	600
DB	541	TGAGCAACACATTGAAATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGATGAAGAT	600
QY	601	ATAGATCCGTATATAATATCCGACGAAATTTGACTCCATTTATTGGAGAAACCGCTAA	660
DB	601	ATAGATCCGTATATAATATCCGACGAAATTTGACTCCATTTATTGGAGAAACCGCTAA	660
QY	661	AATTTCCCTTTGCAATCTCGTCAATCCCTCATTCACCGACCGACCGCATTCGAAACCGTTCG	720
DB	661	AATTTCCCTTTGCAATCTCGTCAATCCCTCATTCACCGACCGACCGCATTCGAAACCGTTCG	720
RESULT 2			
AB029340			
LOCUS		1476 bp mRNA PLN 04-APR-2000	
DEFINITION		Perilla frutescens mRNA for anthocyanin acyltransferase, partial cds.	
ACCESSION		AB029340	
VERSION		AB029340.1 GI:7415645	
KEYWORDS		anthocyanin acyltransferase.	
SOURCE		Perilla frutescens leaf cDNA to mRNA.	
ORGANISM		Perilla frutescens	
		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.	
REFERENCE			
AUTHORS		Sakakibara,K.Y., Tanaka,Y., Mizutani,M.F., Fujiwara,H., Fukui,Y., Ashikari,T., Yamaguchi,M. and Kusumi,T.	
TITLE		Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6	
JOURNAL		Unpublished (1999)	
REFERENCE		2 (bases 1 to 1476)	
AUTHORS		Sakakibara,K.Y. and Tanaka,Y.	
TITLE		Direct Submission	
JOURNAL		Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Keiko Y Sakakibara, Suntory Research Center, Fundamental Research, Plant Biotech, Shimamoto-cho, Wakayamada, 1-1-1, Mishima-gun, Osaka 618-8503, Japan (E-mail:keiko.Sakakibara@suntory.co.jp, Tel:+81-75-962-8807, Fax:+81-75-962-8262)	

```
Qy 721 TTTTCACCAATCCAAATTAAGAAATTAAGAGGTTGGATTGAGTCCAGAGTTCCAAAGTT 780
Db 721 TTTTCACCAATCCAAATTAAGAAATTAAGAGGTTGGATTGAGTCCAGAGTTCCAAAGTT 780
Qy 781 TAGTCACATCTCTCACTTTTGTAGGATGTCAGCTTATATGTGGCTGGCATAACGAAT 840
Db 781 TAGTCACATCTCTCACTTTTGTAGGATGTCAGCTTATATGTGGCTGGCATAACGAAT 840
Qy 841 CATTACAGCAGATGAAGACCAAGACACAGGAGTGCATTTTCTTGATTCGGTTCGATC 900
Db 841 CATTACAGCAGATGAAGACCAAGACACAGGAGTGCATTTTCTTGATTCGGTTCGATC 900
Qy 901 TAAGCCACGATAGATCCGCGGTTCTGAAATTAATCTCGGAACGCTTATCGTAGC 960
Db 901 TAAGCCACGATAGATCCGCGGTTCTGAAATTAATCTCGGAACGCTTATCGTAGC 960
Qy 961 CGCTCCGAGAAATCGGCGGAGAGCTGGTGGGAGAAAGGGTGTCTTGGCAGCTG 1020
Db 961 CGCTCCGAGAAATCGGCGGAGAGCTGGTGGGAGAAAGGGTGTCTTGGCAGCTG 1020
Qy 1021 AGCTAATCGCGCGAGATATAAAAAAGGATCAACGACAAAGAGATATTAGAAACGGTGG 1080
Db 1021 AGCTAATCGCGCGAGATATAAAAAAGGATCAACGACAAAGAGATATTAGAAACGGTGG 1080
Qy 1081 AGAAATGTCGCGGAGATTCGTAAGCGTTGCAGAAATCATATTTTCGTTGGCAGGAT 1140
Db 1081 AGAAATGTCGCGGAGATTCGTAAGCGTTGCAGAAATCATATTTTCGTTGGCAGGAT 1140
Qy 1141 CGAGCAAGCTAGATCTTTACGTTGAGATTTTGGATGGGAGAGCGAGAAAGCAAGAAA 1200
Db 1141 CGAGCAAGCTAGATCTTTACGTTGAGATTTTGGATGGGAGAGCGAGAAAGCAAGAAA 1200
Qy 1201 TATTGTCGATGATGGGAGAAATGCAATGACGCTTTGTAAAGCCAGGATTTTCGAAG 1260
Db 1201 TATTGTCGATGATGGGAGAAATGCAATGACGCTTTGTAAAGCCAGGATTTTCGAAG 1260
Qy 1261 GAGGATGAGGTTTGTCTTTGTCCTTAAGGACAAATGATGATGATGATGATGATGATG 1320
Db 1261 GAGGATGAGGTTTGTCTTTGTCCTTAAGGACAAATGATGATGATGATGATGATGATG 1320
Qy 1321 TTTCACTGGGAATTAATGTTAAATGATGATGATGATGATGATGATGATGATGATGATG 1380
Db 1321 TTTCACTGGGAATTAATGTTAAATGATGATGATGATGATGATGATGATGATGATGATG 1380
Qy 1381 ATTAATTAAGTCTTCAGTAAGTGAAGATTAATCCCTATTATATATATATATATATATAT 1440
Db 1381 ATTAATTAAGTCTTCAGTAAGTGAAGATTAATCCCTATTATATATATATATATATATAT 1440
Qy 1441 AAATAAGTGAAGGCTCTTTGAAAAAAGGAAAAA 1478
Db 1440 AAATAAGT-TAAGGCTCTTTGAAAAAAGGAAAAA 1476

RESULT 3
E12754 1622 bp DNA PAT 24-JUN-1998
LOCUS Gentianatriflora mRNA for acyltransferase, complete cds.
DEFINITION E12754
ACCESSION E12754
VERSION E12754.1 GI:3251586
KEYWORDS JP 1997070290-A/2.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1622)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 2 18-MAR-1997;
SUNTORY LTD
OS Gentianatriflora
COMMENT PN JP 1997070290-A/2
PD 18-MAR-1997
```

```
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12N1:865),
PC (C12N9/10,C12N1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FH source 1..1622
FH /variety='japonica',
FH /organism='Gentiana triflora' FT
FH /tissue_type='petal'
FH /clone='pGAT106'
FT CDS 35..1474
FT /product='acyltransferase'.
FEATURES
source
1..1622
/organism="unidentified"
/db_xref="taxon:32644"
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN
Query Match 14.8%; Score 218.4; DB 10; Length 1622;
Best Local Similarity 50.8%; Pred. No. 9.5e-43;
Matches 742; Conservative 0; Mismatches 661; Indels 57; Gaps 7;
Qy 3 GTGATCGAAAGCTGTAGAGTTGGCGCGCGGACTCGGTGGCGGAGCAATCGGTGGCG 62
Db 62 GTTCTTGAAATGCGGTGTTGGCGCCACCACCGGACGCGTGGCGGAGTTTACAGTCCA 121
Qy 63 CTCACATCTTCGACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCCTCTTCAC 122
Db 122 CTGCTGTTTTCGACATGCGATGGTTGATCTCTGATGAGAACACCATCTGCTATCTAC 181
Qy 123 GAATTCCTCTGTTCCAAAGCAACATTTTTCAGAAATCCATCGTTCACAAACTCAACAATCT 182
Db 182 AGATTCCCGCATCTCTGTTGCCAAC---TCTAAATTTATCATTTTCATCCATTAATTCGTC 238
Qy 183 CTCTCTAAACTCTCATACATCTCTCCCTCTCTCATGCAATTTAATCTACCTTTCA--- 239
Db 239 CTTTCCCTTGTCTCAACACATTTCTTCGGTTAGCGGGAATTTGATTTGGCGGTAGAT 298
Qy 240 TCCCGGAGAAATGCCGAGGTTTCGGTATCTATCCGGGAGCTCGGTTCTTTTCCACCATC 299
Db 299 TCCTCCGATAGATGCCGAGGTTGCGTTTACAGAAAGGGGACTCGGTTTCTTTTAAACAAT 358
Qy 300 GCAGAAATCTAGCGGACTTCGATGATCTCGTTCGGAATTCGTCAGAAATCTCCCTTAGG 359
Db 359 GCAGAAATCGAGATGGATTTGATATCTCGCGGAGATCATCAGAGGATTTCTATATAA 418
Qy 360 CTCTACAACTTTGTCCTTAAATTTGCCCTCATTTGTCGAAGATTCGATAGAAACTCTTC 419
Db 419 TTCAACGATTTGATTCGCGAGCTGCAGAACGATTTGAACCTCGCGGACGACGAATTA 478
Qy 420 CAAGTTTTCGCGGTCAGGTGACTCTTTTCCAGGCCGAGGCTCGGTATTTGGAATAGCA 479
Db 479 CCACTTTGTGTTTACAGGTGACGGTGTCTTCCAAACACCGGTATATGATTTGGACGCAAT 538
Qy 480 ACGCATCACCGTTAGCGAGCGCCCTCGTTTCTCGCGTTTATTAACGGCTTCTTCA 539
Db 539 CTTTATCAAGTCTTGTGGTGGTCCAGTTCTTTCTTCGATTTTAAATTAATGTTGGTTTG 598
Qy 540 ATGACCAACACATTTGAAATGAAGATGAAGATGAAGATTTAAATCTTTGCCAGTTTTC 599
Db 599 GTTGACAAATCCAATGGAGATTCATTAAGTTCCCTTCCACTTTCTCTCTACCTATGAC 658
Qy 600 GATAGATCCGTCATAAAATATCCGAGCAAAATTT-----GACTCCATTTATTTGGAGAAC 653
Db 600 GATAGATCCGTCATAAAATATCCGAGCAAAATTT-----GACTCCATTTATTTGGAGAAC 653
```

```

Db 659 GACAGATCTGTGTCAGAGATCCATTTCATATTCGTGGAATAATCTACAAATGAAGAAAA 718
Qy 654 GCCTAAATATTCCTTTGCAATCTGTC-----ATCCCTCATATTACCGAGCGACCCGAT 707
Db 719 CTGCTCAATCTCAGGGCACACTACTGTTCTAAATCCAGCAATTTCTTAAAGATGAAGTT 778
Qy 708 CGAACACCGTCTGTTTTCACCCCAATCCAAATTAAGAAATTAAGAGGTTGGATTTCAGTCC 767
Db 779 CGAGCCACCTTCATCTCATACACCTATTGATATCATCAAGCTCAAGAAATTCATTTCGTCA 838
Qy 768 AGATTCCAAATTTAGTC-----CATCTCTCATCTTTTGTAGGATTCGA 812
Db 839 AAAATCGCAACTTACCGGTAGTAGTAATAATCTGTCACTTCTACGCTGACATCT 898
Qy 813 GCTTATATGTGGCTGGCATACGAAATCATTC-----ACACGAGATGAAGAC 860
Db 899 GCNCTGATCTGGACATCTGTCGAATCATTTAGACACCGCTGTAAGAGAGAGGTTGGA 958
Qy 861 CAAGCAACGAGGATGATTTTCTTGTATTCGGTGCATCTAAGGCCAGGATTAGATCCG 920
Db 959 GAGGATAAATGCAGCAAACTTATGTCTTCATCACTGCGCAACAGTTTGTCTCG 1018
Qy 921 CCGGTTCCGAAATTAATCTCGGGAATGCTTATCTGACGCGTCCGAGAAATGCGGCGG 980
Db 1019 CCGATACCTCAAAATTTACTTTGGAATTCATAGTGCCTTGTATGTTGGATCGACTCAT 1078
Qy 981 CGAGAGCTGGTGGGAGAGAGGGTCTTCTGGCAGCTGAGTATCTCGGCGGAGATA 1040
Db 1079 GAGCAACTTGTAGGAATGAAGGTTGTGCGTAGCTAGTCAACCGCCATCGGAGATGCTATC 1138
Qy 1041 AAAAAGGATCAACGCAAGAGAAATATTAGAAACGGTGGAGAAATGGTCCGCGGAGATT 1100
Db 1139 CATAAGAGTTACATGACATACGAAAGAAATCTGAGAGAGATTTGGATATCGCGCCCGGA 1198
Qy 1101 CGTAAAGC-----GTTGCAAGAAATCATATTTTTCGGTGGCAGATCGAGCAAG 1148
Db 1199 TCAACATCTCGCGCACCAAGTTCGACGCTCATTTATGCTGTTGGATCCGACACCAAT 1258
Qy 1149 CTAGATCTTTACGGTCGAGATTTTGATGGGGAAGCGGAGCAAGCAAGAAATATTGTG 1208
Db 1259 GTGCATGATTTGATCGAGATTTGTTGGGGAAGCTTGAAAGCATGAATCTGTTTCA 1318
Qy 1209 ATTGATGGGAGAAATATGCAATGACGCTTTGTAAGACGAGGATTTTCAAGAGGAGATTG 1368
Db 1319 ACTAATCTTCGCAACACTAATTTTGTATCTCTCGTCCAGAGATTTAAAGGAGCACT 1378
Qy 1269 GAGGTTGCTGTCTTGGCTAAGGACAAATGGATGCTTTTCTGCTGCTATTTTTCACG 1328
Db 1379 GAGCTTGGCATTTCTTGGCTAAGAAATAGGATGGAGCGCATTTGCCACCATTTTACGAAT 1438
Qy 1329 GGAATTAATGGTTAATAAATGATGTAATTAATACTAATATTATTAATTAACAATTAATTA 1388
Db 1439 TTCAATAGTATCCATGTGAGGAGCCCTTTGTAAGAAAAAGTGTATCAATGATATA 1498
Qy 1389 AGTGTGAGTAACGTGAAGA 1408
Db 1499 AAAAGACAGCAAGATTATGA 1518

RESULT 4
AB026494
LOCUS Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete cds.
DEFINITION
AB026494
VERSION AB026494.1 GI:7415596
KEYWORDS acyltransferase homolog.
SOURCE Gentiana triflora cDNA to mRNA.
ORGANISM Gentiana triflora
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
REFERENCE 1 (sites)

```

```

AUTHORS Yonekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M.,
Fujiwara,H., Fukui,Y., Toshihiko,A., Yamaguchi,M. and Kusumi,T.
TITLE Molecular cloning and biochemical characterization of
hydroxymethyl-CoA:anthocyanin
3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla
frutescens and diverse plant acyltransferase homologs
Unpublished (1999)
REFERENCE 2 (bases 1 to 1622)
AUTHORS Tanaka,Y. and Yonekura-Sakakibara,K.
TITLE Direct Submission
JOURNAL Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Tanaka, Suntory Ltd., Institute for Fundamental Research;
Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan
(E-mail:Yoshikazu.Tanaka@suntory.co.jp, Tel:81-75-962-8807,
Fax:81-75-962-3262)
FEATURES
source Location/Qualifiers
1..1622
/organism="Gentiana triflora"
/db_xref="taxon:55190"
35..1474
/gene="GAT106"
35..1474
/gene="GAT106"
/codon_start=1
/product="acyltransferase homolog"
/db_xref="GI:7415597"
/db_xref="BAA93452.1"
/translation="MAGNSEDIKVLEKRVAPPDAVEFTVPLSFFDMRWLISDEH
HUHYFRHRPNSKFISSIKSYSLVHLKFLPLAGNLIMPVSDSDRMLRYRKGD
SVSLTAIESMDFDYLAGDHQDSYKENDLIPQLEPIVTSGDEVLPALQVTVFSN
TGICIGRNHLQVGLDASSFLHFNKLVLDKNSGSLKFLPLSSLPVMSRVVQDPFH
IRKLYNERKLLKSGTPTVLPALPSKDEVRATFILHPIDIMKLFKFISSKRNLTGS
SNVNLSTFTVTSALITWCLSKSLDTVVREKVEEDKHAANLCAFINCRQFAPIPNY
FNCILVPCNVGSTHEQLVGNELSVNAATGAIDAIHKRLHDYEGILRGDWISPRSTA
APRSTLLIYVGSQARNVDFDADFGWGLEKHEKESVSTNPSTATLILISRRRFGALEL
GISLPKRNMDAFATFTNFINSLHRSPL"
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN

Query Match 14.8%; Score 218.4; DB 12; Length 1622;
Best Local Similarity 50.8%; Pred. No. 9.5e-43;
Matches 742; Conservative 0; Mismatches 661; Indels 57; Gaps 7;
Qy 3 GTCATCGAAACGTTAGAGTTGGCGCGCGCGGACTCGGTGGCGGAGCAATCGGTGCCG 62
Db 62 GTTCTTTGAGAAATAGCGGTGTTGGCGCACCAACCGGCGCGGAGTTTACAGTCCCA 121
Qy 63 CTCACATCTTCGACATGACGTGGCTGCTTTTCATCCCATGCTTCAGCTCCTCTCTAC 122
Db 122 CTGTCGTTTTTCACATGCGATGTTGATCTCTGATCGACAGAACCATCTGCATTTCTAC 181
Qy 123 GAATTCCTTTGTTCCAGCAACATTTTTCAGAAATCCATGTTCCAAACATCAACAATCT 182
Db 182 AGATTCCGCCATCTTGTGCCAAC--TCTAAATTTATCATTTTCATCAATAATCGTCC 238
Qy 183 CTCTCTAAAACTCTCATACACTTCTTCCCTCTCTCATGCAATTTTAATCTACCCCTCA --- 239
Db 239 CTTTCCCTTGTCTCAAAACACTTTTCTCCGTTAGCGCGGAAATTTGATTTGCGGGTAGAT 298
Qy 240 TCCCGGAGAAAATGCGGAGTTTTCGGTATCTATCCGGGACTCGGTTCTTTTCAACCATC 299
Db 299 TCTCCGATAGAAATGCGGAGTTGCGTTACAGAAAGGAGACTCCGTTCTTTTAAACAT 358
Qy 300 GCAGAACTAGCGACACTTTTCATGATCTCGTCGGAATCGTCAGAAATCTCCCGTTAGG 359
Db 359 GCAGAACTAGCGACATGATTTTGATATCTCGCGGAGATCATCAGAGGGATCTTTATAA 418
Qy 360 CTCTACAACTTTCTCCCTAAATTTGTCGAGCAATGTCGAGAAATCCGATAGAAAATCTTC 419
Db 419 TTCAAGGATTTGATTCGCGAGCTGCCAGAACCGATTGTAACTCCGCGCAGCAAGATTGA 478
Qy 420 CAAGTTTTCGCCCTGCGAGGTGACTCTTTTCCAGCGCGGAGCGCTCGGTATTGGAATAGA 479

```

```

Db 479 CCACCTTTTGTCTTACAGGTGACGGTGTCTCCACACCGGTATATGCATTGGACGCAAT 538
Qy 480 ACGCATCACACCGTTAGCGACGCCCGCTGCTTCTCGGTTTATACCGCTTGGTCTTCA 539
Db 539 CTTTCATCAAGTCTTGGTGATGCGAGTCTTCTTTTCGCAATTTTAAATAAATATGCGTTTG 598
Qy 540 ATGAGCAACACATTTGAAATGAAGATGAAGATGAAGATTTAAATCTTTGCGACGTTTC 599
Db 599 GTTGACAAATCCAATGGAGATTCATTAAGTTCCTTCCACTTCTCTACCTATGATAC 658
Qy 600 GATAGATCGGTCAATAAATATCCGACGAAATTT-----GACTCCATTTATTTGGGAAC 653
Db 659 GACAGATCTGGTGCAAGATCCATTTTCATATTCGTGAAATAATCTACAATGAAGAAAA 718
Qy 654 GCGTAAATTTCTTTCGCAATCTGCT-----ATCCCTCATTAACCGACGACCGCAAT 707
Db 719 CTGCTCAAATCTCAGGGCACACCTACTGTTCTAAATCAGCAATTTCTAAAGATGAAGTT 778
Qy 708 CGAACACGTTCTGTTTTCACCCAAATCCAAATTAAGAAATGAAGGTTGGATTCAGTCC 767
Db 779 CGAGCCACTTCATCTACACCCATTCATATCATGAAGTCAAGAAATTCATTTCTGCTCA 838
Qy 768 AGAGTTCCAAGTTTAGTC-----CATCTCTCATCTTTTGTAGCGGATTCGA 812
Db 839 AAAATCGCAACTTAACCGGTAGTAGTAATTAATATATCTGTCAACTTTCACGGTGACATCT 898
Qy 813 GCTTATATGTGGCTGGCATTAACGNAATCATTC-----ACAGCAGATGAACAC 860
Db 899 GCACTGATCTGGACATGCTTTGCGAAATCATTTAGACACCCGTCGTAAGAGAGAAAGGTGAA 958
Qy 861 CAAGACACGAGGATGATCTTTCTTTGATTCGGTTCGATTAAGGCCACGATTAAGATCCG 920
Db 959 GAGGATTAACATCGACGAACTTATGCTCTTTCATCACTGCCGACACGTTTGTCTCCG 1018
Qy 921 CCGGTTCTGAAATTAATCTCGGGAACATGCTTATGTTAGCGGCTGCCGAGATCGCGCG 980
Db 1019 CCGATACCTCAAAATTAATCTTGGAAATGTCATAGTGCCTTGTATGTTGGGATCGACTCAT 1078
Qy 981 CGAGAGCTGGTGGGAGAAAGGGTGTCTTTCGACGCTGAGGTAAATCGCGCGGAGATA 1040
Db 1079 GAGCAACTTGTAGGAAATGAAGGTTGTGCGTAGCTGCAACCGCATCGGAGATGCTATC 1138
Qy 1041 AAAAAAGGATCAACGACAGAGAAATATTAGAAACGGTGGAGAAATGTCGCCGAGAT 1100
Db 1139 CATAAGAGGTTACATGACTACGAAGGATTCAGAGGAGATTTGATATCGCGCCCGCA 1198
Qy 1101 CGTAAAGC-----GTTGCAGAAATCATATTTTTCGTTGGGAGGATCGGACGCAAT 1258
Db 1199 TCAACATCTGCGGCACCAAGGTCACGCTCATTTATGTCGTTGGATCGGACACGCAAT 1258
Qy 1149 CTAGATCTTTACGGTGCAGATTTTGGATGGGGGAGGCGGAGAAACAGAAATATTGTCG 1208
Db 1259 GTGATGATTTTGTATGAGATTTTGGTTGGGAAAGCTTGAAGAGCATGAATCTGTTTCA 1318
Qy 1209 ATTGATGGGAGAAATATGCAATGACGCTTTGTAAAGCCAGGGAATTTGAAAGGAGATTG 1268
Db 1319 ACTAATCCTCGGCAACACTAATTTTGATCTCTCGGTCGCAAGATTTAAGAGGACATT 1378
Qy 1269 GAGGTTGCTGCTTTGCTTAAAGACAAAATGGATGCTTTTGTGCTTATTTTTCAGT 1328
Db 1379 GAGCTTGGCAATTTCTTTCCTTAAAGAAATAGGATGGACGCAATTTGCCACCATTTTACGAAT 1438
Qy 1329 GGAATTAATGCTTAATAATCTATTAATTAATACTAATTAATTAATGTAACAAATTAATA 1398
Db 1439 TTCATCAATAGTCTCCATGTCGAGGAGCCCTTTTGAAGAAAAAAGTGGTATCAATGATAA 1498
Qy 1389 AGTGTGAGTAACTGGAAGA 1408
Db 1499 AAAGACAGACAAGTTATGA 1518
```

RESULT 5

```

E12757 1508 bp DNA PAT 24-JUN-1998
LOCUS Senecio cruentus mRNA for acyltransferase,,partial cds.
DEFINITION E12757
ACCESSION E12757
VERSION E12757.1 GI:3251589
KEYWORDS JP 1997070290-A/5.
SOURCE unidentifed.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1508)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 5 18-MAR-1997;
SUNTORY LTD
COMMENT OS Senecio cruentus
PN JP 1997070290-A/5
PD 18-MAR-1997
PR 30-JAN-1996 JP 1996046534
PF 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAHI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
CC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1508
FT /organism='Senecio cruentus'
FT /tissue_type='petal'
FT /clone='pCAT48'
FT CDS 1..1367
FT /product='acyltransferase'.
FEATURES
source
1..1508
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 442 a 293 c 296 g 477 t
ORIGIN

Query Match 12 98; Score 190.6; DB 10; Length 1508;
Best Local Similarity 49.68; Pred. No. 5.7e-36;
Matches 676; Conservative 0; Mismatches 654; Indels 33; Gaps 6;

Qy 2 CGTGATCGAAACGTTAGAGTTGGCGCCGCCGCGGACATCGGTGGCGAGCAATCGGTGCC 61
Db 5 CATCTCGAACATGCCGAATATCGGCCCTCGGGCACCATCGGCCATCGCTCGTTATC 64
Qy 62 GCTCACATTTTCGACATGAGTGGTGCATTTTCATCCCATGCTTCAGTCTCTTCTTA 121
Db 65 TCTTACTTTCTTCGACATTAATTTGCTACTCTTCCCTCGGTGCCACCATCTTTCTTCTA 124
Qy 122 CGAATTCCTTTGTCACAGCAACATTTTTCAGATCCATGTCCTCAAAACACCAACAATC 181
Db 125 TGACTTTCCACATCTAAATCCCATTTTCATGGACACTATTTGCCAGGCTAAAAACAATC 184
Qy 182 TCTCTTAAACACTCTCATACACATCTTCTCCCTCTCTCATGCAATTTAAATCTACCCCTTCATC 241
Db 185 TTTATCGTCACTCTTCACATTTTTCCTCGTTTTCCTAGTAAATTTGATTTGATTTCTTAA 244
Qy 242 CCCGG-----AGAAATCCCGAGTTTCGGTATCTATCCGGGAGTCCGT 286
Db 245 CACTGATGGTTTGGGTTTAAATAAAAAACAGAAATAAAACACGTTGAAGGTGATCTCTGT 304
Qy 287 TTTCTTACCATCGCAATCTAGCGAGACTTCGATCATCTCGTCGGAATTCGTCAGAC 346
Db 305 TGTGGTTACTTTTCGAGAAATGTTGCTTGAATTTTAAATAATTTGACAGGAAATCATCTCGC 364
Qy 347 ATCTCCCGTTAGGCTCTACAACTTTTGTCCCTAAATTTGCCGCCCAATTTGTCGAAGAATCCGA 406
```


Db	365	AAATGTGAAACTTTTATCCACTTGTACCTTCAATGGGAAATGCAATCAAAATATATGTGA	424
Qy	407	TAGAAAACTCTTCCAAAGTTTTCGCCCTGCAGGTGACTCTTTTCCAGGCGGAGCGCTCGG	466
Db	425	TTCCGTCAGCGTCCCACTTTTTCACCTCAAGAGAGCGTTTTCGGGGCTCGGGTATATC	484
Qy	467	TATGTGAATAGCAACGATCACACCGCTTAGCGACGCCCGCTCGTCTCCGCTTTTATAAC	526
Db	485	ACTAGGAATGACGAATCATATAGCCTTGGTGACGCTAGCACGCGTTCAACTTTTGGAA	544
Qy	527	GGCTTGGCTTCAANTGAGCAACACATTTGAAATATGAAGATGAAGATGAAGAATTTAAAT	586
Db	545	AGGTGGACTTCGATTTATTCATCTGCTGTAGATCGCTCTTTTAAACGAAAGG---ATC	601
Qy	587	TTTGGCAGTTTTCGATAGATCCGCTCAATAAATATCCGACGAATTTGACTCCATTTATG	646
Db	602	TCACCGGTTTTTGTAGATGAT---TGATTAACATCCACATTTAGATGAATAAGTTGAG	658
Qy	647	GAGAAACGGGCTAAAATTTCTTTGCAATCTCGTATCTCCCTCATACCGACGACGCGCAT	706
Db	659	ACATACAGGCTCGAAGTTTTTATAACCTTCGAGCCTTGTGGTCCCACTGATAAGT	718
Qy	707	TGGAACACGCTTGGTTTTCACCAATCCAAATTAAGAAATTAAGAGTTGGATTCAGTC	766
Db	719	TCGGTCAACGCTTGTGTGACCGCAACTAATATCAATCTACTAAAGAAAAGGCTCTTAAC	778
Qy	767	CAGATGTCGAAGTTTGTAGTCCATCTCTCATCTTTTGTAGCGATTCGACGTTATATGGGC	826
Db	779	CCAAGTGCCNAACCTTGGGTACATGTCATCTTTTACGGTAACCTTGTGGTTATATATGAG	838
Qy	827	TGCATTAACGAAATCATTC-----ACAGCAGATGAAGACCAAGCAACAGGAGTCATT	880
Db	839	TTGCATAGGAAATCACTCGTAAATATAGGAGAAGAAAGGGCGAGACGAGTTAGAAC	898
Qy	881	TTTCTGATTCGGTGCATTAAGGCCACGATTAGATTCGCCCGGTTCTCGTAAATTAATT	940
Db	899	GTTTCATAATCACCATTGATGTGCTGCTCTGCTGTGATCCACCAATTCACACGCTACT	958
Qy	941	CGGAACGCTTATCTGACCGGTCGCGAGAAATCGCGCGGAGACGCTGTGGGAGAGAA	1000
Db	959	TGTAACATGTTGGTGACCACTGTCTCCGACCTTAAAAAATGTCTGTTTGAAGCGGAAA	1018
Qy	1001	AGGGGTGTTTCTGGCAGCTGAGGTAATCGCGCGGAGATAAAAAAGAGTCAACGACAA	1060
Db	1019	TGGGTATGCACCTTGGTGCTAAAGTAATTTGGAGAGTCTATATGCAAAATGATATATA	1078
Qy	1061	GAGAAATTTAGAAACGGTGAGAAATGGTCCCGGAGATTCGTAAGCGTTGCAGAAATC	1120
Db	1079	GGACGGAATCTTGAAGATGCCCGAGATGGCATGA---ACCTTTTCATGATCCCGGCTAG	1135
Qy	1121	ATATTTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTACGGTGCAGATTTTCGATGGG	1180
Db	1136	GAAGATTTGGTGTGCTGGTACACCTAAGCTCAACTTGTACGACTTTGATTTTGGGTGGG	1195
Qy	1181	GAAGCGGAGAAACGAGAAATATTTGCGATTTGATGGGAGAAATATGCAATGACGCTTG	1240
Db	1196	GAAGCCGATAAAGATGAGACTGTTTCAATAGA---CTAATAACGTCGATTTCTATAAA	1252
Qy	1241	TAAAGCCAGGATTTTCAAGAGGAGATTTGGAGGTTGCTTCTCTTTGCTTAAGGACAAAT	1300
Db	1253	TGCAAGCAAAACATCAGCACAAGATCTTGAAATTTGATTTGAGTCTACCGAGTATGCAAT	1312
Qy	1301	GGATGCTTTTGCCTATTTTTCAGTGGGAATTAATGGTTAA	1343
Db	1313	GGAGCGGTTTTCAGCATCTTTGATGAAGGATTAGAGAGTCAA	1355

RESULT 6
AB010708
LOCUS AB010708 1679 bp mRNA PLN 20-FEB-1999
DEFINITION Gentiana triflora mRNA for Anthocyanin 5-aromatic acyltransferase,
complete cds.

ACCESSION	AB010708	AB010708.1	GI:4185598
VERSION			
KEYWORDS			Anthocyanin 5-aromatic acyltransferase.
SOURCE			Gentiana triflora petal cDNA to mRNA, clone: pGAT4.
ORGANISM			Gentiana triflora
REFERENCE			Eukaryota; Virdiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
AUTHORS			Fujiwara, H., Tanaka, Y. and Kusumi, T.
TITLE			Direct Submission
JOURNAL			Submitted (22-JAN-1998) to the DDBJ/EMBL/GenBank databases.
MEDLINE			Hiroiyuki Fujiwara, Suntary LTD., Inst. of Fundamental Research; 1-1-1, Wakayamadai, Shamamoto-cho, Osaka 618-0024, Japan
FEATURES			(E-mail: Hiroiyuki.Fujiwara@suntary.co.jp, Tel.: +81-75-962-8807, Fax: +81-75-962-8262)
Source			2 (sites)
CDS			Fujiwara, H., Tanaka, Y., Yonekura-Sakakibara, K., Fukuchi-Mizutani, M., Nakao, M., Fukui, Y., Yamaguchi, M., Ashikari, T. and Kusumi, T. cDNA cloning, gene expression and subcellular localization of anthocyanin 5-aromatic acyltransferase from Gentiana triflora Plant J. 16 (4): 421-431 (1998)
polyA_site			Location/Qualifiers 1..1679 /organism="Gentiana triflora" /db_xref="taxon:55190" /clone="pGAT4" /tissue_type="petal" 6..1415 /codon_start=1 /product="Anthocyanin 5-aromatic acyltransferase" /protein_id="BAA74428.1" /db_xref="GI:4185598"
BASE COUNT	488	a	350 c 359 g 482 t
ORIGIN			1679 /note="34 a nucleotides"
Query Match			11.6%; Score 172.2; DB 12; Length 1679;
Best Local Similarity			49.5%; Pred. No. 1.8e-31;
Matches	701;	Conservative	0; Mismatches 668; Indels 48; Gaps 8;
Qy	48	GAGCAATCGGCTGCACATCTTTCGACATGACGCTGGCTGCATTTTCATCCCATCCTT	107
Db	78	GAGTTATCGCTACCGGTAACATCTTCGATATCCCTGGTTGCACCTGAATAGATCGAG	137
Qy	108	CAGTCTCTTTCAGCAATTCCTTGTTCACAGCAACATTTTTCAGCAATTCAGTTCCTCA	167
Db	138	TCCTTCTGTTTACGACTTTCGCTACCCCAAGAACACATTTCTTTGACACTGTTATCCCT	197
Qy	168	AAACTCAAAACATCTCTCTTAAACCTCTCATACATCTTCCCTCTCTCATCAATTTA	227
Db	198	ATCTTAAGGCCCTCTTGTCTCTCATCTCTAAACACATACGCTTCGCTTACGGGAATTC	257
Qy	228	ATCTACCTTTCATCCCGGAGAAATGCCGAGTTTTCGGTA-----TCTATCCGGGAC	281
Db	258	TTGATGCCGATCAATCGCGGAAATGCCGAGTTTCAGTACTCCCGCTGATGAGGGCGAC	317
Qy	282	TCGGTTTCTTTCACGATCGGCAAGATCTAGGACGACTTCGATGATCTCGTCGGAATCGT	341
Db	318	TCGATAACTTTGATCGTTGCGGACTCTGACGAGGATTTTGACTACCTTAAAGGTCATCAA	377

Qy	1386	TTAAGTGTGTTAGTACGTGGAAGATAATAATCCCTATAT	1422
Db	1452	CCAAGTCCCTTTAGTAACACTACACCAACCCCTACTTT	1488
7			
RESULT	7		
E12753			
LOCUS	E12753	1703 bp	DNA
DEFINITION	Gentianatriflora mRNA	acetyltransferase, complete cds.	24-JUN-1998
ACCESSION	E12753		
VERSION	E12753.1	GI:3251585	
KEYWORDS	JP 1997070290-A/1.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1703)		
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.		
TITLE	GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY		
JOURNAL	Patent: JP 1997070290-A 1 18-MAR-1997;		
COMMENT	SUNTORY LTD		
	OS Gentianatriflora		
	PN JP 1997070290-A/1		
	PD 18-MAR-1997		
	PF 30-JAN-1996 JP 1996046534		
	PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915		PI
	ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHIRO,		
	PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAAKI		PC
	C12N15/09, A01H1/00, C07K14/04, C07K14/42, C12N9/10, (C12N9/10, PC		
	C12R1:865)		
	PC (C12N9/10, C12R1:19);		
	CC strandedness: Double;		
	CC topology: Linear;		
	FH Key	Location/Qualifiers	
	FH		
	FT	1. .1703	
	FT	/organism='Gentianatriflora'	
	FT	/variety='japonica'	
	FT	/tissue_type='petal'	
	FT	/clone='pGAT4'	
	FT	6. .1415	
	FT	/product='acyltransferase precursor'	FT
	mat_peptide	9. .1412	
	FT	Location/Qualifiers	
		1. .1703	
		/organism='unidentified'	
		/db_xref='taxon:32644'	
BASE COUNT	512 a	350 c	359 g 482 t
ORIGIN			
Query Match			
Best Local Similarity 11.6%; Score 172.2; DB 10; Length 1703;			
Matches 701; Conservative 0; Mismatches 668; Indels 48; Gaps			
Qy	48	GAGCAATCGGTGCGGTACATCTTCGACATGACGTGGCTGCATTTTCATCCCATGCTT	107
Db	78	GAGTATCGGTACCGGTAACATCTTCGATATCCCTGGTTGCATTTGAATAAGATGCAG	137
Qy	108	CAGTCTCTTCTACGAATTCCTTGTTTCCAAGCAACATTTTCAGAAATCCATCGTTTCCA	167
Db	138	TCCCTCTCTGTTTACGACTTTCGGTACCCGAACACATTTCTTGGACATCGTTATCCCT	197
Qy	168	AAATCAAAACAATCTCTCTCTAAACATCTCATACATCTTCTCCCTCTCTCATGCAATTTA	227
Db	198	AATCTTAAGGCCCTCTTGTCTCTCACTCTAAACATACGTTCCGCTTACGGGAATTTG	257
Qy	228	ATCTACCTTCAATCCCGGAGAAAATCGCGAGTTTCGGTA-----TCTATCGGGGAC	281
Db	258	TTGATGCGGATCAAAATCGGGGAAATCCGGAAGTTTTCAGTACTCCCGGTGATGAGGGGAC	317

```
QY 282 TCGGTTCTTTTCCACATCCGAGAACTAGCGAGCTTCGATGATCTCGTCGGAATCGT 341
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 TCGATAACTTTTGATCGTTCCGGAGCTGACCGAGGATTTTGACTACCTTAAAGGTCATCA 377
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 342 CCAGAAATCCCGTTAGGCTCTACAACTTTTGCCCTAAATTCGCGGCCAATTTGTCGAAGAA 401
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 CTGGTAGATTCCAATGATTGATGGCCCTTTTATGTTATGCCAGGGGTTATAAGGACC 437
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 402 T---CCGATAGAAAACCTCTTCCAAGTTTTCGCGCTGCGTCACTCTTTTCCAGGCCCA 458
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 ATGCAAGACTATAAGTGATCCGCTGAGCGTGCAAGTAAACCGTTTCTTCTTAACCGT 497
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 459 GCGCTCGGTTATGGAAATAGCAACGATCACACCGTTTAGCGAGCCCGCTGTTTCTCGCG 518
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 GGCATAGCCGTGCTCTGAGCGCACATCAATCAATGCGAGATGCTATAAAGTTTGTGAATG 557
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 519 TTTATAACCGCTTGGCTTCTCAATGAGCAACACATTTGAAATGAAAGATGAAGTGAAGAA 578
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 TTTCATCAATGCTTGGGCCCTATATTACAAA---TTTGGGAAAAGACGCGACTTGTGTGCC 614
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 579 TTTAAATCTTTGCCAGTTTTCGATAGATCCGTCATAAAATATCCGACGAAATTTGACTCC 638
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 615 GCGAATCTTCTTCCATCTTTCGATAGATCGAATCAAAAGATCTGTATGGCCTAGAGGAA 674
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 639 ATTTATTGGAGAAACGCGCTAAATTTTCCT-----TTGCAATCTCGTCATCCCTCA 689
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 675 ACATTTTGGACGAATGCAAGATGTTCTTGAATGTTCTTAGATTTTGAAGCAACACC 734
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 690 TTACCGACGCGCATTCGAACGAGTTGCTTTTTCACCCCAATCCAAATTAAGAAATTG 749
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 735 CCTCGATTCAACAAGGTAGGAGCTACATATGTCCTCTCCCTTGTGAAATCCAGAGCTA 794
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 750 AAG-----GGTTGGATTCACTCAGTCCAGATTCCAAGTTTAGTCCATCTCTCATCTTTGTA 803
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 795 AAGAACAAGTACTGAATCTCAGAGATCCGAACCCGACAAATACGTGTACGAGCTTTCACA 854
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 804 GCGATTGCGAGCTTATATGTGGGCTGCATAACGAATCATTCACAGCAGATGAAG----- 858
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 855 ATGACGTGTGGATACGTATGGACATGCATGGTCAAAATCAAAAGATGACGTCTATCAGAG 914
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 859 -----ACAAGACAACGAGGATGATTTTTCGATTCCGGTTCGATCTAAGGCCA 908
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 915 GAATCATCAACGACGCAAAATAGCTCGAGTACTTTCAGTTTACACGCGGATTCGCCGAGGA 974
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 909 CGATTAGATCCCGCTTCTCGAAATTAATCTCGGGAATCTGCTTATCTACGCGCTGCCG 968
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 975 CTCTGACCCCGCTGCTCGGCTAACTACTTTTGGCAACTGCTTCGCGTACGCGTTGCA 1034
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 969 AGAATCGCGCGGAGAGCTGGTGAGAGAAAGGGGTGTTTCTGCGAGCTGAGGTAATC 1028
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1035 AAAGCAACACATAAAGAGTTAGTTGGGATAAAGGCTTCTTGTTCAGTTGCAGCTATT 1094
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1029 GCGCGGAGATATAAAAAAGGAT---CAACGACAGAGATAATTAGAACGGTGGAGAA 1085
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1095 GGAGAAAGCCATTGAAAAGAGTTGCAACAACGAAAAGGGGTTCTTCAGATGCAAAAACCT 1154
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1086 TGTGTCGCCGAGATTCGTAAGCGTTGCGAAGATCATATTTTTCGTCGCGCAGGATCGAGC 1145
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1155 TGGTTATCGGAATCTAATGGAATCCCTTCAAAAGATTTTCTCGGGATTACCGGATCGCCT 1214
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1146 AAGCTAGATCTTTACGGTGCAGATTTTGGATGGGGAAGCGGAGAAAGCAAGAAATATTG 1205
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1215 AAGTTCGATTTCGTATGCTAGATTTTGGATGGGAAAGCCCTGCAAAATTTGCATTACC 1274
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1206 TCGATTGTTGGGAGAAATATGCAATGACGCTTTGTAAAGCCAGGATTTTCGAAGGAGGA 1265
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1275 TCTGTTGATTATGCAAAATG---ATTTATGTAATTCAGTCCAGGATTTTGAAAAAGGT 1331
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1266 TTGGAGGTTTGTCTTTCCTTAGGACAAAATGGATGCTTTTCTGCTCTATTTTCA 1325
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1332 GTGGAGATTGGAGTATCATTTGCTTAGATTCATATGATGCTATTTGCAAAAAATCTTTGAA 1391
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1326 CTGGGAATTAATGGTTAATAAATGTATGTAATTAATAATATTATTATGTAACAAATTAA 1385
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
Db 1392 GAAGGCTTTTGCTCTTTGTCATAGTCTCTTTAATAGAACCATATTTGCTGCAATAAAGTA 1451
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 1386 TTAAGTCTTGAGTACGCTCAAGATAAATCCCTATTAT 1422
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1452 CCAAGTCTCTTATAGTAACTACACCAACACCTACTTT 1488
||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
LOCUS AB016892 66237 bp DNA PLN 27-DEC-2000
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MXF12.
ACCESSION AB016892 BA000015
VERSION AB016892.1 GI:3449333
KEYWORDS Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
SOURCE Clone:MXF12.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (sites)
AUTHORS Asamizu,E., Sato,S., Kaneko,T., Nakamura,Y., Kotani,H., Miyajima,N.
and Tabata,S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 5. VIII.
Sequence features of the regions of 1,081,958 bp covered by
seventeen physically assigned P1 and TAC clones
JOURNAL DNA Res. 5 (6), 379-391 (1998)
MEDLINE 99156233
REFERENCE 2 (bases 1 to 66237)
AUTHORS Nakamura,Y.
TITLE Direct Submission
JOURNAL Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935; Fax:81-438-52-3934)
COMMENT Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MXF12
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremling1.cool.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is K15E6 and the 3' clone is K3K3.

FEATURES             Location/Qualifiers
     1..56237
         /organism="Arabidopsis thaliana"
         /strain="Columbia"
         /db_xref="taxon:3702"
         /chromosome="5"
         /clone="MXF12"
         /clone_lib="Mitsui P1"
         /note="gene_id:MXF12.2"
         /codon_start=1
         /evidence=not_experimental
         /product="receptor-like protein kinase"
         /protein_id="BAB10824.1"
         /db_xref="GI:1017545"
         /translation="MIRHALIFSLVSTPIVGEGATSTYPTDVFLENCGDTSNNDV
VSGRNWTAERHOKILSSNLVNASFTAAQVQESGVQIPYMTARIPSEFTYSFPVTPG

CDS
```

```
SNFLRLYFYPTRGQNAVKSFFSVKVNQGTLLNNSADLTWKASKPQTEFIKEFI
IPVQTLNLFTSLDSLAFNVEIGTIVSNREYSKGGFDVITNNGSSVDFHIESVA
AFETVRLNVTGDSGMFRWRWVSDEIILSESSGISIPVDIKINITEKTPSYST
PDDVYATSRMGNADHPQONLNLWLTFTVDAGFSYLRLHFCETLSSEVNGEGRVF
SIFTEHOTATDNDVPMGSGSWIPMYLDYTVTAGSGSRHDLRLDHLPLSNPKY
YDALLNGVEILKNDPNCNLAGNPDLVSPDLIPNRPATRIKKNSHIIPITLAVVG
SLVAMFVGVVLIMKKKSKPSTNSWCPLPHGTDSTNTPAKSLPADLCRRFSI
FEKSAFNVEDKLIIGVGGFSGYKQIDGGATLVAVKRLAITSNQAKFEFTELEM
LSKRLNHLVSLIGCYCEDENMLVLYEYMPHGTLKDLHFRDKTSOPLSWKRRLEIC
IGAARGLOYLHTGAKYTI IHRDITKNILLDENFNTKVSDFGLSRVSGPTASOTHVST
VVKGTCFLDPEYRROVLTEKSDVYSFGVVLVLELCCRPIMQSVPEQADLIRWVK
SNVRGIVDQIISDLSADITSLEKFCIEIARCVCQDQRGKPPMDVNVVWALEFALQ
LHETAKKNDNVSZLSDMGSEVGTITTDGDDLLFSRTTGHVGKSTTTDSDVVLVVGDR
SGSGWGYFSINEPKAR"
6701..7210
/notes="gene_id:MXF12.3
/known protein"
/codon_start=1
/evidence=not_experimental
/protein_id="BAB10825.1"
/db_xref="GI:10177546"
/translacion="MINLEKEAKEKOILLRDKRKIVHHVNTLLVKIYEASALDORH
LFLKPYGPNYFIKAWLDDTKEYGYVVKDFKVNQPPKIPLNPPAOYLIVELLK
GMSTRDPGTNRVTVMGRAKIRUSPLMTSYKERPISGGIDLVLGNSDQCVVKKGYLY
AMKLHRSAD"
7670..10111
/notes="gene_id:MXF12.4"
/codon_start=1
/evidence=not_experimental
/product="receptor protein kinase-like protein"
/protein_id="BAB10826.1"
/db_xref="GI:10177547"
/translacion="MNCNVLFSLVLSVVTAGVTAAVHPTDVFLENCGDTSNNVDNSG
RNNTVRSQIILSNLNAVSTSEASQKAGVSRIPYMKARIESEFTYSFPVTPGSLF
LRLYFYQTKSGFDVANSFVSVKYNGFTLLRNFNADSTVQASIPLSNLIKEFIIPV
HQTLNFTGTSKNLLAEVNGIEIYVMPDRFYSKGGFDNLRNVSSDVFQDINDSTAFE
VYATSRMLGNSNLMFNLTGFTLDAGYNLVLRLHFCETLPQVTRAGQRFVSIFVED
KMAKKTVDRLSGGPRIPIYLDVSVYVFGESGMIOQLDLRLDPLKDTNQTYYDAIL
SGVEILKNDSDNLRNPPELLAVSTDDTDDSNVTPPIKGGPHVLVILIVVGSVIG
LATFIVIMLLIRQMKRKNKENSVMFKLLLKQYIAELKKITISFSTHTVKGKGG
TVVRGNSNGRTVAVKVLKQNGDDFINETVMSQTSQSHVNVLSLGLFCYEGSKRAI
ISEFLNGLDQVQISRNKSTPNVTLDYIALGIARGLEYLHGCRTRIYVHFDPQN
ILDDNFCPVQDFGLAKCEKRESILSLDITRTGITYIAPEVSRMYGISHKSDYV
SYGMLVDIMGARKVETTCNGSTAYFPDWIYKOLENGDQWTIGDEINEEDNKIVK
KMLVSLWCIRPCPSDRPPNKKVVMIEGSLDALELPPKPSRHSITELVLESLSLDG
QEAKEQTQILDSTII"
10819..13239
/notes="gene_id:MXF12.5"
/codon_start=1
/evidence=not_experimental
/product="receptor protein kinase-like protein"
/protein_id="BAB10827.1"
/db_xref="GI:10177548"
/translacion="MICFILFVSFLVSATAPYKPDVDFLINCGETVFPDNGHRT
WTOEKNILPKNSDNASFSVWSYKESGIPQVPMYMTARIFRSDTYTSPFVSPGKWF
RLYFYPTYSKSGFDVANSFVSVNDFTLLQNEADLTVKASIPESKSLIKETIPVY
LTLNLRPNSNLAFTVNGIEIYVMPDRFYSKGGFDLITNVGSLDIEDNSTASET
VHRLNVGHNVDESDGMFRRLWSDDEYFELIGVSGMPDPMVNLSTFEKTPAYAVAP
VYTCRMMQNAQDTYI LNLFNLWLTFTVDAGFSYLVRHFEKYLANKANQRFVSIFLG
NOMAREMDYIRLSGGRPIYILDFTRIYVGGESGPRDLRLDHLPLVKONPEYFEAIL
NGVEILKLNNSGNLAIQDNEKPNPLSNVTQOIKGKSSHLVKIFITAVGP
GTGLATPVVVLMLWMQMRKRNKEERVNFKLLNMTYAEULKKITKFSYLIIGKG
FCVYVGNLSNGRKVAVKLDLKGSAEDFINEVASMSQTSQSHVNVLSLGLFCFEGSKR
ATVYFELNGLSQFMRNKSQTQDVITLIGALGTARGLEYLHYGCKTRTIYHFIDKP
QNILLDNLCPKVSDFGLAKCEKRESILSLDITRTGITYIAPEVSRMYGRVSHSKD
VYSFGMLVDIMGARKVETTCNGSTAYFPDWIYKOLENGDQWTIGDEINEEDNKIEK
EIAKMLVGLWCITQPCPSDRPPNKKVVMIEGSLDALEIIPKPSRHSITELVLES
LSDGEDV"
join(16412..16498,16859..16924,17003..17038,17085..17192,
17337..17461,17541..17640,17803..17974,18073..18191,
18285..18395,18486..18632,18712..18844,18934..19009,
19232..19297,19385..19492,19591..19698,19873..20070,
```

```
20164..20373)
/notes="gene_id:MXF12.6"
/codon_start=1
/evidence=not_experimental
/product="ABC transporter-like protein"
/protein_id="BAB10828.1"
/db_xref="GI:10177549"
/translacion="MGNKKLITGSSKTHGSSSYRDLQNOEDKPKANGSENGLD
LEHGVAAVNAVNGFRFALICQYISIMVSOAKPDAGLVITGTTALLIGSTNTLLVPKF
GGMIDIVSRDVKTPQOETSLIARNVAVIILLVVGISICTALRAWLNSASERVV
SLGDKCPVGNPDGVDLNDVWFVSPRSPSHMLKIGLSRLTSGKVALGVSGGKKT
IANLIERFYDPLAGKILLNGVSLMEISHOYLHKQISIVSOEPIILFNCVSENIAYGPD
GEASFTDIENAKMANAHEIFAPPKYNTVVGERGLRLSGGQKQRIARALLTNPS
VLILLDEATSALDAESEYLVQDAMDSDMAGRTVILVIAHRLSTVTKTCAVADSVDSGEVAE
KGTHDELLSLNGIYTNLKRLOLQSSSSVTTL"
join(25348..26359,26401..26756)
/notes="gene_id:MXF12.7"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin acyltransferase-like protein"
/protein_id="BAB10829.1"
/db_xref="GI:10177550"
/translacion="MVNEEMESSLKVIDVARVTPSNSDSESSTLPLTFDOLLWKYLH
AVERVIFYKLTDASRPFDSDVIVPNLKTSLSSLSHYLPLAGKLIWEPLOPKIYVT
PNDVSTVAESNADFSRLGKEPFTTEIPLVPELHVSDSASAVSQVTLFPNQG
FCISVNAHVAVDGKTTNFKLKWARTCKNQDSFLPDQLIPVDRDTVTKDPMDLDTKI
LNAMHRVAKYFTGGKEPNKSLKLLMSPEIGDPVFRYTLNTREDIQKRLERLKKES
SSSSSSSPKELRLSTFVIYVSYALTCLIKARGDPSRPVGYGFVDCRSLMVPVPS
SYPTPMSEGEFLAAARMVSDVEALDENVALKIPEILEGFTTLLSPCTQVLSVAGSTR
FGYVGDLPDQGRPEKVVVVSIDOGAELSPAESRDGSGVGLFSLKKHEMDVLVDLLH
KGLN"
complement(join(27466..27879,28046..29073))
/notes="gene_id:MXF12.8"
/codon_start=1
/pseudo
/evidence=not_experimental
/product="transposase-like protein"
complement(join(30804..30874,30947..31289))
/notes="contains similarity to Ac-like transposase"
gene_id:MXF12.9"
/codon_start=1
/pseudo
/evidence=not_experimental
32432..33823
/notes="gene_id:MXF12.10"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin acyltransferase-like protein"
/protein_id="BAB10830.1"
/db_xref="GI:10177551"
/translacion="MVSSLNIEVAVRTPSNSDASAELSPLTYFDLIYKLRVERV
IFRTVNTVPFFDSVIVPNLKTSLSSLSHYLPLAGKLIWEPLOPKIYVTIYSONDV
SFSAVFNADFSLSLGNPEPSTELVPLALOSSDASITVSQVTLFQNOECIGV
SFAHVLADGKTTTFLKSWAHCKHQDFSLQDLPYDRTVTKSPDTSNKVLEWR
SFTKILAGGCEPANKSLKLNPSFEIGDPVVRVTLQTRDITOTLERLKRREVVSSSS
STSSKELRLSTFVIYVYLVIIRARGGEHPRVGYAFSDVCRSLMNPPTNYFGN
CIAGSRMLLTAKMFGEGLAAATMVSDSIBEWDESFAWKIPDFVAYATLPPETOL
ILVSGSNRFQVYELDFGWRGPRDKVMVVSIPGNGISMAESRDQNGSVEIGFSLKKHEM
DTLIDLHCELT"
34363..35709
Query Match 7.5%; Score 110.8; DB 12; Length 66237;
Best Local Similarity 46.5%; Pred. No. 3.1e-16;
Matches 616; Conservative 0; Mismatches 662; Indels 48; Gaps 6;
```

```
Qy 37 ACTCGGTGGCGGACAACTCGGTCCGCTCACATCTTCGACATGACGCTGCATTTTC 96
||||| ||||| | | | ||||| ||||| ||||| ||||| |
Db 34418 ACTCGTGGCGGTCACTCACTCTCCGCTCACTTTCTTTGACCTCTCTGCTCAACACA 34477
```


JOURNAL	Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE	7 (bases 1 to 101176)
AUTHORS	Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C., Khan S., Kim C., Altafi H., Bel B., Chin C., Chlou J., Choi E., Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M., Palm C., Pam P., Sakano H., Schwartz J., Southwick A., Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A. and Ecker J.
TITLE	Direct Submission
JOURNAL	Submitted (19-JUL-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
COMMENT	On Jul 15, 2000 this sequence version replaced gi:2618677.
FEATURES	Location/Qualifiers
source	1..101176 /organism="Arabidopsis thaliana" /db_xref="taxon:3702" /chromosome="1" /clone="F21B7"
CDS	join(736..1038,1276..1713,1874..2119,2205..3032,3117..3274,3483..3610,3725..4134,4313..4435) /note="unknown protein; similar to ESTs dbj AV529515.1, dbj AV529266.1, dbj AV554950.1, dbj AV523761.1, dbj AV523761.1, dbj AI996703.1, and dbj AV546743.1" /codon_start=1 /evidence=not_experimental /product="F21B7.1" /protein_id="AAF86508.1" /db_xref="GI:9280639"
CDS	/translation="MQPLPARSGDHEGFWNSHPLLVAAGDTNGTGLGHFSONGSL ARDGSKAGDATNYPTTFVYSLRSHSYVVLRFSSVCMIRCSRVAVGLANGI YCDALALEKHSVLTPVPQVRQTRVNVGYPMAVGRWLYASKSMKTGR LDPSTFSPSLSPSSSGSSGMARYAMESKOLANGLNGLMGDKYTKLSKYCDML PDGSTPASNAIKGVGSCDAENAGAVKDLVSGALVSQFKAHTSKYLCDFP SGTLTVASVCGNNINVFQIMPSRSHNAPGDLSEWESSHVHLFKLHRGITSAIVDI CFQSOQWVAIISKGGCHIFVLNSSGSDAFOPECEETRLPASLPPWFQTSLSS NQSLSPATAVLSVRIKYSFSGWLTNVSNTAATGKVFVPSGAAVAFKSVTH DLQNSTNALEHLVLTIPSGHVQHELLFSVCTESPENGLRVTOKTSHVQVQDDLRV KVEPQMDWDCRLEETELPKSITEKOYDLETNVSNHLTSHEDACLDMNSHFS EDYKSCSKPERSHCYLSNFVKYVTSGLPWNQSKISFHVMDSPRSSSTGGGF EIEKVAHELEIKQKLLPVDFHFKATLEDKRFMSKCYHTSATGSHQVNGKICQDI INCHSKPGSIESASEEGSTKOMENLHSDHMSNISKSLPLYPTVNGIYKETEKN ANGMKPVTAKLSTLAKETRTNGFTTPIILTOSVNEQMLSTGCKPMPGFGFALHEHC KAVADPEHLKKLDEVTNVHNLNVNNNTEKLGQDKMVSQVLAFAFGCILYKDFL RVGNFNSVPFLLDYKILWLRK" complement(4838..6223) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F21B7.2" /protein_id="AAF86541.1" /db_xref="GI:9280672"
CDS	/translation="MASCIELHFTHLHIPVTINQOFLVHPSSPTPANQSPHLSLYLS NLDDICARVETSVSYEPTNNRESVLKRLDALSEVLVPPVPLSGRLREVENGLK EVFGEQGLVMTYSANSSMDLADLTVPNPAWPLIFRPGEEAYKILEMPLLIAQ VTFPTCGGFSGLJRLCHICIDCGFCAQMOFLGSWAATAKGLIADPEVMDRETFKPN PPWKYPHHIEVLPERSLNTNLDWTKPLQKCYRISKFCQCRVKSIAQGEDPTLVCS TFDAAAHIRWSWALDKVPLDPLNDRLTFESVNVTRLETLLKRLKKGYGNVCLACAM SVSESLNLSLSTRLVQDARLRVSDYVDVVRKPRLEFGGKLTITQWTR FTRLLSLDNHK" complement(join(6700..7357,7482..7884,8620..8839,8920..9241,9758..10380)) /note="unknown protein; similar to ESTs gb T04168.1, gb T1964.1, and gb AW004462.1" /codon_start=1 /evidence=not_experimental /product="F21B7.3" /protein_id="AAF86540.1" /db_xref="GI:9280671"
CDS	/translation="MGHDSFCYLIVLRALRCGIIAQICALQKERRSKMSSDRS SQADEFETGVKGLVASGKIPEAMFHTPPDTLTSLKOTAPPSQOLTIPTVDLKG SMDLISRSVETKIGDAAEKFFOVVNHGISEVVMERKGEIRRFHEQDQDEVKRFY SRDTRDVLVYSNIDLHTCNKAANRDTLACYMAPDPKQLQDLPAVCGETIMWEYSKOL MTLGEFPELLSEALGNPNHLKDKGCAKSHIMFGQYPPCPKDFLIGTSKHRTDSF ITILLQNIQGLQVHDCQWVDSVPFALVLDLLQLISNDKFIISAEHRJANGS SEPRISCFVSTEMKPNRIYGIPELLSEQNPAKYRDLTIEFSNTFSQTIIRTE IPALFRATPATLASLKPPPKHLTIPTVDLKASVAVKEIGEAEEKGLFHLVNHGIP VEPLERMIQIRGFHEOEPAKRFYSRDRDVLVFSNHDQNSEAKSARDTLCGYT AVERPREDLPVAVCGEIMLVEYSKEIMSLGERFELLSEALGNSHLKDKMDCAKSOYM VGQHYPCPQDPLTIGINKHTDISFLVLLQDNVGGLOVHEQWIDVTPVPGALVIN IGDFLQKIMLLRSCFISCMIWGLLILFYQVLITNDKFIISAEHRVIANGSEBPT SVAIVFTFMRAYSRVYGPINKDLSAENPAKYRDLTIEFSNTFSSKTLDPKLLHHEK I" 10833..11348 /note="unknown protein; similar to EST emb Z30724.1" /codon_start=1 /evidence=not_experimental /product="F21B7.4" /protein_id="AAF86509.1" /db_xref="GI:9280640"
CDS	/translation="MRGFSSVSSWLTSTVEIGEVFLGSTWLDVAVAEDAVDHWVES ACLMVNLKDLRWLLRVPVAREGLSEPFPSLRFPSIRIEGCFAGKFSDLRWRSE AMKCVVLRWGLDQIGWRSVGNWVSVRFVAGPASSLVRDLRQDSYTSANDLPEVVV LQSSGPCIRRY" join(111965..12033,12173..12303,12448..12519,12596..12685,12756..12915,12985..13041) /note="similar to ATHP1 dbj BAA37110.1" /codon_start=1 /evidence=not_experimental /product="F21B7.5" /protein_id="AAF86510.1" /db_xref="GI:9280641"
CDS	/translation="MNTIVVAQLQRQDQIVISLYQOQFLDNQFSELKRLQDEGTPDF VAEVSVUFFDDCSKLINTMSLERPDNVDFKQVDGSHQLKGSSSLCFLVLSGARR VKNVCISFEKCCDQVNRREGYSSNSHTNCDNDIWHRCRLCQQVDYKMLKTKLQDLF NVSKFLCENLVLLELLKQILQAGGTIPQVDIN" 16495..17682 /note="similar to protein kinase pir A57676; similar to ESTs gb A1993651.1, dbj AV538995.1, dbj AV522524.1, and dbj AV538995.1" /codon_start=1 /evidence=not_experimental /product="F21B7.6" /protein_id="AAF86511.1" /db_xref="GI:9280642"
CDS	/translation="MRETKLVWCLMFLLRGFTTEALIDPVDFLALQAIKRLSDLLPG SKFESWDFTSDFGAGVYCNQDKVLSNLGDDPRAGSLGRIDIPAIKLSALTTEL SIVPGRIMGALPATISQKDLRFLAISRNFIISGEPASLGAEVGRGLRTLDLSYNQLTQ ISPSIGSLPELSNLIILCHNHTGSIPLPQLTIRIDLKRNSITGSIPLSPSLSQY LSLANOLTGSVYHVLRLNQLNYLDLSLNRTGTIPARVFAFPTNLQLRNFFGL IQPANQVYITSDLSYNRFSGGISPLSSVENLYLSNRFTGEVPASFEVERLLSANIQ TLXLQHNFLTGIQISPAAEIPVSSSLCQVNCMVPPLQTPCLKAGQKTRPTTQCTE WRG" complement(join(18462..18515,18612..18742,18840..18975,19064..19139,19214..19410,19497..19595,19668..19752,19865..20103,20179..20253,20345..20649,20798..20902,21088..21216,21356..21437,21513..21608,21780..21849,21929..22025,22408..22502,22586..22659,22810..22888,23041..23101,23206..23315,23411..23556)) /note="hypothetical protein" /codon_start=1 /evidence=not_experimental /product="F21B7.7" /protein_id="AAF86539.1" /db_xref="GI:9280670"

Query Match	7.5%	Score 110.8	DB 12	Length 101176
Best Local Similarity	47.2%	Pred. No. 3.3e-16		
Matches 651	Conservative	0	Mismatches 692	Indels 35
Gaps				
Qy	2	CGTCATCGAAACGNGTAGAGTTGGGCGCGCCGCGACTCGGTGGCGGAGCAATCGGTGCC	61	
Db				
Qy	37273	CATCATCGAGACCTGTGCACATCTCTCCCCCAAGGGACCGGTGCATCCACCACCTCTTCC	37332	
Db				
Qy	62	GCTCACATCTTCTTGACATGAGCTGGCTGCATTTTCATCTCCCATCGTTCAGCTTCCTTCTCTA	121	
Db				
Qy	37333	TCTCACCTTCTTCGATGCCCTCGCTCTCTCTCCACTCGCCGATTCCTCTCTTCTTCTT	37392	
Db				
Qy	122	CGAATTCCTTGTTCGAAGCAACATTTTTTCAGAAATCGATCGTTCCAAACATCAACAATC	181	
Db				
Qy	37393	CTCTTAGCAAACTCAACTGAATCTTCTCTCAAGACTTTGTAGCCAACCTCAACAATTC	37452	
Db				
Qy	182	TCTCTCTAAAACCTCTCATACACTTCTCCCTCTCTCATGCAAAATTAATCTACCCCTTCATC	241	
Db				
Qy	37453	CCTCTCCATCACTCTCAACATTTCTTCCCTTAGCCCGGTAACTGATTTCCCGCCTCG	37512	
Db				
Qy	242	CCCGGAGAAAATGCCGGATTTTCGGTATCTATCCGGGGAAGCTCGGTTCCTTCACCATCGC	301	
Db				
Qy	37513	TCCCGACCTCCATATTTACACTCAACGCTGGCGAAGACTCTCTGTTTTCCACGTAGC	37572	
Db				
Qy	302	AGAATCTAGCGA----CGACTTCGATGATCTCGTCGGAAATCGTCAGAAATCTCCGTTAG	358	
Db				
Qy	37573	AGAGTCTACTGAAACCGCACTTCGACCAACTGAAGTCTGATTTCCCTTAAAGATATCAGTGT	37632	
Db				
Qy	359	GCTCTACAACTTTCTCCCTAAATTTCCGCCCAATTTGTCGAAGAATCCGATAGAA---AACT	415	
Db				
Qy	37633	GTTGCATGGCGTCTTGCCCAAGTTTACCTCTCTTCCACTCGCTCTCCGGAAGAAATTCAGAT	37692	
Db				
Qy	416	CTTCCAAGTTTTCGCCGTGAGTGACTCTTTTCCAGCGCGAGCGTGCGTATTGGAAT	475	
Db				
Qy	37693	CGGACCAATATGGCAATGCAAGTCAACATTTCCCTGGAGCTGGAATCTGTATAGGCA	37752	
Db				
Qy	476	AGCAACGCATCACACCGTTAGCGAGCGCCCGTCTTCGCGTTTATAACGGCTTGGTC	535	
Db				
Qy	37753	CTCAGCTACACATGTTGTAGCAGATGGAGTCACCTTCAGTCATTTTCATGAAGTATTGGAT	37812	
Db				
Qy	536	TTCAATGAGCAACACATTTGAAATGAAGATGAAGATGAAGAATTTAACTCTTTGC----	591	
Db				
Qy	37813	GTCAITGACCAATTCACGCGCAAGATTCGCGCACGGTTCTTCTACCTCTCTGCGCTAT	37872	
Db				
Qy	592	---CAGTTTTCGATAGATCCGTCATAAAATATCCGACGAAATTTGACTTCCATTTATTGGA	648	
Db				
Qy	37873	TCACAGCTGCAGAACATAATCAAGGCCAGGAGAGGTACCCGACGACATTTGGAGCG	37932	
Db				
Qy	649	GAACGGCGTAAAATTTCCCTTTGCAATCTCGTCATCCCTCAATACCGACGACCGCATTC	708	
Db				
Qy	37933	GTTTTGGAGCCAAACTCTGCAAAACACAGTTTCATGTTACGCGTGA-GAACAATGGTCA	37991	
Db				
Qy	709	GAACACGTTCTGTTTCCACCAATCCAAAATTAAGAAATTCAGGGTTGGATTCAGTCCA	768	
Db				
Qy	37992	GAGCCACTTTTACATTTAGCCGGAAGCAGATAGACAATCTCAAGAGTTGGGTTCACAGAGC	38051	
Db				
Qy	769	GAGTTCCAAGTTTAGTCCATCTCTCATCTTTGTAGCGATTGCAGCTTATATGTGGGGCTG	828	
Db				
Qy	38052	AGTCGAGAATCAATCTCTCTTCTACCTTCGTGGTGACTCTAGCTTTCATTTGGGTAA	38111	
Db				
Qy	829	GCATAACGAAATCATTCACAGCATGAAGACCAAGACACAGGAGTGCATTT-----TT	882	
Db				
Qy	38112	GCTTGATTAAACACACTTGTTCAGACAGTGAACAGAGGCCCAAGCAAGTCTTTTCACT	38171	
Db				
Qy	883	TCTT		

	I I
Dn	38292 AAAAAATGCTCATGTGCTGCTTCAGATGAACATCACAGCAGAGAATCAAGATGATTGTATCAA 38351
Qy	1057 ACAACAGAAATATTACAACCGTGGGAGAAATGCTCCCGGAGATTCTGTAAGCGTTGCAGA 1116
Dn	38352 CGGA---TCTGTTGAAGACACACCAGATGGGACAGGACTACGTAATGGGTATGT 38408
Qy	1117 AATCATATT---TTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTACGGTGCGAGATTTTG 1173
Dn	38409 CCCATTATCCAACCTCAATTGCAGCAGCTCCGAAATTCGGACTGTATGACATGGATTTTG 38468
Qy	1174 GATGGGGAAGCGGAGAAAACNAGAAAATAATYCTTCGATTGATGGGAGAAAATATGCAATGA 1233
Dn	38469 GATTAGGAAAGCCCTGCAAAATGAGGATAGTGCACATAGAAACAG---GTGGTTCTATCG 38525
Qy	1234 CGCTTTGTAAAGCCAGGATTTTCGAAGGAGGATTCGAGGTTTGTCTTTTGCCTTAAGG 1293
Dn	38526 CATTCTCAGATCCAGACAGCGCCACCAATGGAGTTGAGATTCGAATAGCACTAGAGAAGA 38585
Qy	1294 ACAAAATGGATGCTTTTTCGCTTATTTTTCCTACTCGGAATTAATGTTTAATAAATGTA 1351
Dn	38586 AGAAAAATGGATGTGTTTGACTTCAC'TTACAAAAGGGGATCAAGAAATTCGCAACGTA 38643
RESULT 10	
AC003027	AC003027 119914 bp DNA PLN 30-JAN-1999
LOCUS	Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,
DEFINITION	complete sequence.
ACCESSION	AC003027
VERSION	AC003027.1 GI:4079614
KEYWORDS	HTG.
SOURCE	thale cress.
ORGANISM	Arabidopsis thaliana
REFERENCE	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
AUTHORS	Magnoliophyta; eudicotyledons; core eucots; Rosidae; eurosids II;
	Brassicales; Brassicaceae; Arabidopsis.
	1 (bases 1 to 119914)
	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
	Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
	Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
	Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
	Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Unpublished
REFERENCE	2 (bases 1 to 119914)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
	Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,
	Oj,L.O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vyotskaia,V.,
	Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE	Direct Submission
JOURNAL	Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA
	Sequencing and Technology Center, 855 California Avenue, Palo Alto,
	CA 94304, USA
REFERENCE	3 (bases 1 to 119914)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
	Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
	Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
	Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
	Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Direct Submission
TITLE	Submitted (30-DEC-1998) DNA Sequencing and Technology Center,
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304,
	USA
REFERENCE	4 (bases 1 to 119914)
AUTHORS	Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
	Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
	Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
	Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vyotskaia,V.,
	Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL	Direct Submission
TITLE	Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
JOURNAL	Stanford University, 855 California Avenue, Palo Alto, CA 94304,
	USA

COMMENT

On Dec 30, 1998 this sequence version replaced gi:2734094.
 Bases 1-9262 of clone F21M11 overlap with bases 68998-78259 of
 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone
 F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7
 (AC002560).

e-mail for correspondence: arab@sequence.stanford.edu
 Genes with similarity to proteins in the databases are described as
 'putative', '-like' or 'similar to'. Genes that have EST
 similarity but no significant protein similarity are described as
 'unknown proteins'. Genes that are annotated based only on gene
 prediction software are described as 'hypothetical proteins'.
 The software programs used to predict genes include: Grail
 (Informatics Group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge,
<http://genomic.stanford.edu/~chris/GENSCANW.html>), Fexa (V Solovyev
 & A Salanov, Sanger Centre, <http://genomic.sanger.ac.uk/>), and
 NetPlantGene (http://www.cbs.dtu.dk/NetPlantGene.html).
 Denmark, (S.W. Hebsgaard, et al., CBS, Technical University of
 Denmark, Copenhagen).

FEATURES

source

1. 119914

/organism="Arabidopsis thaliana"

/cultivar="Columbia"

/db_xref="taxon:3702"

/chromosome="1"

/clone="F21M11"

1. 7000

misc_feature

/note="Proteins in this region are annotated in the F21M11

entry, AC002411."

misc_feature

1. 9262

/note="Overlap with bases 68998-78259 of 'IGF' BAC clone

F21M11."

gene

7684..10913

/gene="F21M11.1"

join(7684..7866,7965..8027,8122..8160,8250..8381,

8458..8728,8875..9893,10003,10256,10344..10410,

10501..10551,10685..10781,10885..10913)

/gene="F21M11.1"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAD10665.1"

/db_xref="GI:4204284"

/translation="MAPNLRKACDAMKLLGISETKYTRAFRLKLLTYENNWDPIEE
 DAYKVLDAIFDEARDAOSTEKNEKKEKRSVATSGRRKKAPELVPDDEE
 DMDDEDFPLRRGRSSSSSSSYNNEDLTKQPEEDEDGVTLPPLKRY
 VRRGERGLAMTVNNAAPSRLSEPEVPPMVLPAHPMTKYSEASALVILND
 EPNTDHPVSDTNCAPMLEMGKSNHYQEWDETKDILNDTAMDVSPSSAIGES
 SEHKVAASVELASTSGEAKICLSEAPATGETNLHLPMSMEDLRAMEKCLKSYKI
 VHPFESVLGPMKMCSCYIDLAKNSQLLETVCMDSRAGDESGAVGISMLVVPV
 ECEISDGWKAISNKDITAGEENVEIPWVNEINEKVPSPRYMPSHFVQDAPVIFS
 LSPFSDQSCSTIEDCLASEMNCAGVNDNGFAYTLDLGLKEEFLEARISARDQ
 RKQVRFCECPLEKRAKVEILEPCGHLKRGAKCEWFCGCTKCGNRVVGGMHN
 KLOVFFTPNGKMGWGLRTLEKLPKGAFCICEYIGETLITPELYQSFEDKPTLPVILDAH
 WGSERLEGKALCLDGMFYGNISRFNLHRCLDANLIEIPVQVETPOHYHLAFTT
 RDIAMELAWDYGDIDNDMLMKPFCDCGSCFCNRKRSKRTMQLNKA"
 join(13033..13488,13717..13926,14139..14514)
 /gene="F21M11.2"

mRNA

13033..14514

/gene="F21M11.2"

join(13061..13488,13717..13926,14139..14316)

/gene="F21M11.2"

/note="Similar to acid phosphatase; Location of ESTs

110C2T7, gb|T42036, and 110C2XP, gb|A1100245"

/codon_start=1

/protein_id="AAD10666.1"

/db_xref="GI:4204285"

gene

/translation="MDRTMFLSLTIALLVGVVSAGDWNILNQLRGLGSSSQNGIVS
 RKITDLKGCESNRINVEHNIRKDFVPOECVSHIKDYMTSSQYKDDVARTVDEVI
 LHFGMCCSKCDMDAIFDIDITLLSPYHKHKGFGGKLNKFGEDWTQKKK
 QVPHPMKKLYHDIRERGIKIFLISRKEYLRSATVDNLIOAGYIGNSMLRGLDQ
 AKVQYKSKERKWLMSLGYRVMGVMDGQSSFAGCPPLPRRTFKLPNSIYVVA"
 complement(join(14004..14312,14565..14641,14787..15831,
 15906..16006,16327..16362))

mRNA

15906..16006

/gene="F21M11.3"

gene

complement(14004..16362)

/gene="F21M11.3"

complement(join(14567..14641,14787..15831,15906..16006,

16327..16362))

/gene="F21M11.3"

/note="Unknown protein; Location of ESTs 40C3TT,

gb|AA728590 and 40C3T7, gb|T04573"

/codon_start=1

/protein_id="AAD10667.1"

/db_xref="GI:4204286"

/translation="MLDQISGNTHKLSVETAEPHLNDRVHIVEIPKASVPIPIE
 ICDAEEKSPSTISRRKRVFDSKVITYEHVSVSESEVELSEENEESEKRSUKSSK
 TDQIIIEVANSNGSYNHRKNCRESDDIDEEDFDCSDLDDEEYYSVDFGKE
 DLHNPTKEVYTDIGDKTEIDSKLRSNETVRDGNHYDQGVNLNENVTOWKSAK
 SKRTKOKQSKENSPIADQEEKRSDSSFEQTDIDITLSVKPKCRIEBPKLNRNOE
 LAVDASLSTLSTSESSECSNASMYLTPEKLSKTSYCSYKPLRINHDDRPVLCALTL
 IDKQFSAVTPRKSPDEPTPICTGGYMGNRKAIDCGSASSFKGIPNTSSKY
 REDKSNVHSTPFPEARLEKALNNIDK"
 complement(16835..20238)

/gene="F21M11.4"

complement(join(16835..17185,17274..17392,17491..17656,

17795..17885,17982..18079,18175..18361,18504..18604,

18705..19049,19134..19349,19439..19791,19862..19967,

20188..20238))

/gene="F21M11.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAD10668.1"

/db_xref="GI:4204288"

/translation="MGSSSPPEARAAQVPSMILFLEICTVHVHYNRRKLNDRVLSA
 NLNPKRVRDLRFMESVSIYKSLNAAVSVHLPQIPNDCYNKDALNNSPKH
 GESESDMTDKVSKRGSDTSSSDGSPLTSESDPRPKHQDWTKQLSDHLLYE
 FESEYDAANHPTPESTYEQAAKNVNDITASEQPSNAARKICGDSFIOESSPNPKTQD
 TLLRLMESLRSDOPTDYVKAQNHOMVYSEFVVIYSFVLCRLSNLRGVVLCVQOLPK
 SHFEQDSKRKRDIRITASDAMENHLKVPKRENNLQKADIDCNGKCSANSDDLSEKIS
 KALEQTSNITICGFCQSAVSEATGEMLHYSRGPVDDGDDIFRSNVIHVSACIEWA
 POVYEGDTVKNLKAELARGMKIKCTCKSLKGAALGCFVKSCRRSYHVPCEATISRCR
 WYEDFLLCPAHSSVFPNEKSGHRSRAEPLKINPAELCSLEQTPAFTKEVLGKW
 SALSCKDKLMESLAVRFNATISRYWNPSVTHVIASTDEKGACTRTKLVLGILNGKW
 IINAAWMSKLSAQPVDERPFEIQIDTQCGDQPKTARLRAETNKPKLEGLKPYFF
 GDFYKGYKEDLONLVKAGCTILNTEDELGAESNNVNDORSSIVYINIDPPHGICAL
 GEVYTIWQRANDREALAQTGSRVLGHTWVLESIAGYKLPKHPVIG"
 20752..24647

/gene="F21M11.5"

join(20752..20994,21362..21497,21596..21740,21825..21994,

22102..22178,22542..22758,22920..23064,23343..23442,

23599..23693,24021..24072,24227..24298,24441..24647)

/gene="F21M11.5"

/note="Hypothetical protein"

/codon_start=1

/protein_id="AAD10668.1"

/db_xref="GI:4204287"

/translation="MGANSKSVTASFVIAVFFLLCGRTAVEDTEFEHGDYSKLSGI
 IIPFASTQLRAWSLDCPYTPDNDLVLDTTKLLSAVNCWFKCMVLDVYNQYD
 HPCKSRPDSGLSAITELDFGYITGLPSTYMKWKEWVEFGEIENAIIVAPYDMRLS
 PTKLERPDYFHLKLTFTALKLRGSPISVFAHSMGNVFRYLEWLRLEIAPKHYL
 KWLDDHAYFVAGPLGSAIKSTYLSGVTGCLPVSEGTARLLSSFASSILWMPF
 SKNCKGDNTEFTHESGGAACKRVYHCDDEEYQSKYSGMPTNIIETIPSTARELA
 DGLFLKAIEDYDPSKRMHLQKLYVFFVIRNIAHRSSLAGFLLYHDDPVFNPLTFW
 ERPPKRVFCIYGAHLKTEVGYFAPSGKPYPDNWIITDIITEGSLSVSRSGTVDDG
 NAGPTIGDETVPVHSWCKNWLGPKNITMAPQIILKIQKQHPDEGSDVHVELNVD
 HEHGSIDIANNTRAPRVKITYTFEDESSEIPGKRTAVWELDKSGY"
 27777..28734

/gene="F21M11.6"

27777..28734

/gene="F21M11.6"

28007..28465

/gene="F21M11.6"

/note="Unknown protein; Location of ESTs 203I24T7,
 gb|H76794 and 203I24XP, gb|AA605510"
 /codon_start=1
 /protein_id="AAD10670.1"
 /db_xref="GI:4204289"

/translation="NMTKMLRPPRVLTADKKRDERDAFISSVTDNPPPIAKFPSPFP
KLPVPVNPISKSTAAAEPSQNLMLAGYLSHEYLTGTLFGQMNQARAQESS
KIKPSHTVPEACECPKRRRYREVANLLRSDAQLPGIWPNAQLARFLK"
complement(29264..32033)
/gene="P21M11.7"
complement(join(29264..31015,31312..31414,31484..31587,
31782..32033))

Query Match 6.7%; Score 99.6; DB 12; Length 119914;
Best Local Similarity 47.3%; Pred. No. 1.8e-13;
Matches 657; Conservative 0; Mismatches 684; Indels 47; Gaps 10;
QY 4 TGATCGAAGCGTAGAGTTGGCCGCCGCCGCGACTCGGTGGCGGAGCAATCGGTGCCGC 63
Db 45957 TCATCGAGACCTGCACATCTCTCCCAAGGCGACCGTTCCATCAACCACCTCTTCCCTC 46016
QY 64 TCACATCTTCGACATGACGTGCGTGCATTTTCATCCCATGCTTCAGCTCCTCTTCTACG 123
Db 46017 TCACCTTCTTCGATGCCCTCGCTCTCTCCCACTCGCCGATTTCTCTCTTCTTCTCT 46076
QY 124 AATTCCCTTGTTCAGCAACATTTTTCAGAAATCCATCGTTCCAAATCTCAAAACATCAATCTC 183
Db 46077 CTTTACCAAACTCACTGAATCTTCTCTCCAGACTTTGTACCAACCTCAAAACATTTCCC 46136
QY 184 TCCTCTAAACTCTCATACACTTCTCCCTCTCTCATGCAATTTAATCTACCTTTCATCCC 243
Db 46137 TCCTCATCTACTCTCAACATTTCTTCCCTTACCGCGTAAATGATTAATCCCGCTCGTC 46196
QY 244 CGAGAAATGCCGGAGTTTCGGTATCTATCCGGGACTCGGTTTCTTCCACCATCGCAG 303
Db 46197 CCGACCTCCATATTTACACTACAGCATGGCCAAAGACTCTCTTGTTTTCCACCGTAGCAG 46256
QY 304 AATCTAGCGA---CGACTTCGATGATCTCGTCGGAATCTCCAGAACTCTCCGTTAGGC 360
Db 46257 AGTCTACTGAACCGACTTCGACCACTGAAGTCTGATTTCCCTTAAGATATCAGTGTGT 46316
QY 361 TCTACAACTTTGTCCCTAAATTTGCCGCCCATTTGTCGAAGAAATCCGATAGAA---AACTCT 417
Db 46317 TGCATGGCGTCTTGCCCAAGTTACCTCTCTCACGCTCTCTCCGGAAGAAATTCAGATGC 46376
QY 418 TCCAAGTTTTCGGCGTGCAGGTGACTCTTTTCCAGCCGAGCGGTGCGGTATTTGGAAATAG 477
Db 46377 GACCAATATGGGTATGCAAGTCAACATCTTCCCTGGAGCTGGAATCTGTATAGGCAACT 46436
QY 478 CAACGCATCACACCGTTAGCGACGCCCGCTGCTTCTCGGCTTTTATAACGCTTTGGTCTT 537
Db 46437 CAGCTACACATGTTGAGCAGTGGAGTCACTTTCAGTCAATTTTCATGAAGTATTTGGATGT 46496
QY 538 CAATGAGCAACACATTTGAAATGAAGATGAAGATGAAGAAATTTAAATCTTTTCG----- 591
Db 46497 CATTGACCAATCCAGCGTTAAGATCCCGCCACGGTTCTTCTACCTCTCTCGCTATTTC 46556
QY 592 -CAGTTTTCGATAGATCGGTCAATAATATCCGACGAAATTTGACTCCCATTTATTTGGAGA 650
Db 46557 ACAGCTGCAGAAACATGATCAAGGACCCAGCGAGGTAGCGCAGGACATTTTGAGCGGT 46616
QY 651 AACCGCTTAAATTTCTTTTGAATCTCGTCATCCCTCATTTACGACGCGACCCGATTCGA 710
Db 46617 TTTGGAGCCCAAACTCTGCAAAACACAGTTTCACATGTTAGCCCTGA-GAACATGGTCAGA 46675
QY 711 ACCACGTTCTGTTTTCACCCCAATCCAAATTTAAGAAATTTGAAGGTTTGGATTTACGTCAGA 770
Db 46676 GCCACITTTTACATTTGAGCCGAGCAGATAGATAATCTAAAGTTTGGGTTACAGAGCAG 46735
QY 771 GTTCCAAAGTTAGTCATCTCTCATCTTTTGTAGCGAATTCGACGCTTATATGTGGCGTGGC 830
Db 46736 TCTGAGAATCAATCTCTGTTTCTACCTTCTGTTGCTGACTCTAGCTTTTCATTTGGGTAAGC 46795
QY 831 ATACGAAATCATTT-----CACAGCAGATCAACGACCAAGCAACAGCAGAGAT 875
Db 46796 TTGATCAAGACACTTTGTACAAAGACAGTGAACCAAGCCCAACGAGGAACAGAGATGAA 46855
QY 876 GCATTT---TTCCTTGTATTCCTGGTCTGATCTAAGGCCACGATTAGA-----TCCGCCCGTT 926

Db 46856 GTATTTCACTTGTATGATCTATGTTGAGTAATCGCCTCAAGTACACACAACCAATA 46915
QY 927 CCTGAAATTTACTTTCGGAACTGCTTATCTAGCGCTCGGAGAAATCGCGCGGAGAG 986
Db 46916 CCACAACATCATTTTGGCAACTGTATGGCTTCCTGGTATCCTATCATGATCAAGAAACACGAT 46975
QY 987 CTGGTGGGAGAGAAAGGGTGTCTTCTGGCAGCTCAGTAATCGCGCGGAGATAAAAAA 1046
Db 46976 TTGCTAGGAGAAAAATCGCTTTTGGCGCTTCAGATGCAATCACAGGAGAAATCAAGAT 47035
QY 1047 AGGATCAACGACAGAGATATTTAGAAACCGTGGAGAAATGGTCCGCGGAGATTCGTAAA 1106
Db 47036 ATGTTATCAAGCGA---TCTGTTCAAGACAGCACCAAGATGGGACAAAGGAGTACGTAAA 47092
QY 1107 GCCTTCAGAAATCATAT---TTTTCGGTGCAGGATCGAGCAAGCTAGATCTTTACGGT 1163
Db 47093 TGGGTCATGTCCTCATTTCAACCTCAATTTGCAGGAGCTCCGAAATTTGGGAGTGTATGAC 47152
QY 1164 GCAGATTTTGGGAGGAGGAGGAGAAAGCAAGAAATATTTGCGATTTGATGGGAGAAA 1223
Db 47153 ATGATTTTGGATAGGAAGCCCTGCAAAATGGAGATCTGCACATAGAAACAG---GT 47209
QY 1224 TATGCAATGACGCTTTGTAAAGCCAGGATTTTCAAGAGGAGATTTGGAGTTTGGCTTCT 1283
Db 47210 GGTCTATCGCATTTCTCAGAGTCCAGAGACGCGCAATGGAGTTGAGATTTGGAATAGCA 47269
QY 1284 TTGCCTAAGCACAAATGGATGCTTTTGTCTTCTTATTTTCACTGGGAATTAATGGTTAA 1343
Db 47270 CTAGAGAAGAAATTTGGATGTTTGTACTCAATTTCTACAACAGGGGATCAAGAAATTC 47329
QY 1344 TAAATGTA 1351
Db 47330 GCAACGTA 47337
RESULT 11
AX025514
LOCUS AX025514 1648 bp DNA PAT 16-SEP-2000
DEFINITION Sequence 40 from Patent WO0032789.
ACCESSION AX025514
VERSION AX025514.1 GI:10187184
KEYWORDS
SOURCE Citrus limon.
ORGANISM Citrus limon
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Sapindales; Rutaceae; Citrus.
REFERENCE 1 (bases 1 to 1648)
AUTHORS Aharoni,A., Verhoeven,H.A., Luecker,J., O'Connell,A.P. and Van
Tunen,A.J.
TITLE Fruit flavour related genes and use thereof
JOURNAL Patent: WO 0032789-A 40 08-JUN-2000;
AHARONI ASAPH (IL); VERHOEVEN HARRIE ADRIANUS (NL); LUECKER JOOST
(NL); CPRO DLO (NL); CONNELL ANN PATRICIA O (NL); TUNEN ARJEN
JOHANNES VAN (NL)
FEATURES
source Location/Qualifiers
1..1648
/organism="Citrus limon"
/db_xref="taxon:2708"
52..1527
/note="cdna
Lemon acyl transferase"
/codon_start=1
/protein_id="CAC09066.1"
/db_xref="GI:10187185"
CDS
52..1527
/note="cdna
Lemon acyl transferase"
/codon_start=1
/protein_id="CAC09066.1"
/db_xref="GI:10187185"
NPPVRLFFVEMTDLTWDLNSELPLKXLSFTLLHYLPLAGHLMWPLDAKPAVY
YFPDNDGVSEVAEWSSECHAGFHLSCNGIRQAVEFHPPLQSLTDDKAEVTAIQ
IYLPNQGESIGVSSHAILDCKSTLPLKSNAYLCKQLQCHHPCLSPELPLDRT
VILFDPGQDMLQLNKVVGDSDNPQIRSLKVLFPDSESLNKLPVATFEUTREDIT
KURHKVNHQLSKSSKQVRLSTFVLTLAYFVCMAKAKLAKAKAEAAAGDETKN
IIVGFTADYRSRLDPPILPNYFGNCNGRHCETAKASDFVQENGVAFAEMLSDVWKG

DADAIEANDKVBILKKEGAMIFSVAGSTQFDVYGSDFGWRPKKVEIVSIDRTQ
AISAARRDGGGVEGVVLEKQOMEFESVFADGLKNDLV"
BASE COUNT 464 a 344 c 354 g 486 t
ORIGIN

Query Match 5.2%; Score 76.8; DB 9; Length 1648;
Best Local Similarity 50.9%; Pred. No. 3.4e-08;
Matches 342; Conservative 0; Mismatches 312; Indels 18; Gaps 6;
QY 812 AGCTTATATGTGGCTGGCATACCAACCAATCATTCACAGCATGAAGACCAAGACAAGA 871
DB 990 AGCTAAATAGCAAAAGCAAACTGAAGCTGAAGCTGCAGCAGGTAATGATGAATAA 1049
QY 872 GGATGCATTTTCTTGTATTCGGCTGATCTAAGGCCACGATTTAGATCCGCCGTTTCCTGA 931
DB 1050 AAATATTATGTGGGATTCAGTCGCGATATATAGGACCGTTTGGATCTCTCAATCCACT 1109
QY 932 AAATTACTTCGGGAACCTGCTTATCTAGCGCTGCCGGAATCGCGCGCGAGACTCGT 991
DB 1110 TAATTATTTCGTAACCTGC--AATGGAGACATTTGTGAGACTGCACAAAGCAAGTGATTTC 1167
QY 992 GGGAGAGAAGGGTGTCTGCGACGCTGAGTAACTCCGGCGGAGATAAAAAAGGAT 1051
DB 1168 GTCAAGAAATGGGGTGTCT--TTTGTTCAGAGATGTTAAGTGATATGGTCAAAAGGAT 1226
QY 1052 CAACGACAGAGAATATTAGAAACGGTGGAGAAATGGTCGCGGAGATTCGTAAGCGTT 1111
DB 1227 CGATG--CGGATGCCATTGAAGCCAATGATGATAAGTTTCAGAAATATTGGAATCTG 1284
QY 1112 GCAGAA-----ATCATATTTTTCGTGGCAGATCGAGCAAGCTAGACTTTTACGGTGCA 1166
DB 1285 AAGAAGGAGCAATGATTTTCTGTGCTGCTGCACCCCAATTTGATGTTTACGGTGC 1344
QY 1167 GATTTTGGATGGGGAAGCGGAGAAAGCAAGCAATATTGTCGATTGATGGGAGAAATAT 1226
DB 1345 GATTTTGGGTTGGGAGGCCCAAGAAAGTGGAGATGTTGTCATATAGTAG--GACACAA 1401
QY 1227 GCAATGACGCTTTGTAAGCCAGGAGATTGCAAGGAGGATGGAGGTTTGTGTCCTTG 1286
DB 1402 GCATCTCTTTGGCAGAGAGAGATGGAGGCGCGTGGAGTTGGAGTCTTTTA 1461
QY 1287 CTAAGGACAAATGATGCTTTTCTGCTCTATTTTTCACATGGGAATTAATGGTTAATA 1346
DB 1462 GAGAAGCAAAATGCAGGTTTTCGAATCTGATTTGCTGCTGAGACTGAA-----AAATG 1516
QY 1347 ATGATGTAATTAACATAATATTATATGTAACAAATTAATTAAGTGTGAGTAACGTGAA 1406
DB 1517 ATCTGTTTAATTAATGATGATCATCATCAATTTCTCAATATATTATTTGGTCATATCAA 1576
QY 1407 GAATAATCCCTATTATATATTATGATTTGGTTCAATAAAGTGTAAAGCCTCTTTGAAAA 1466
DB 1577 AAGAAATAATATTTCGGGATTTTGTGACCACCAATAAATACTCTTTTGTGAAAAA 1636
QY 1467 AAAAAAATAA 1478
DB 1637 AAAAAAATAA 1648

RESULT 12
AB028618 85690 bp DNA PLN 27-DEC-2000
LOCUS Arabidopsis thaliana genomic DNA, chromosome 3, pl clone:MOD1.
DEFINITION Arabidopsis thaliana
ACCESSION AB028618 BA000014
VERSION AB028618.1 GI:5041971
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (sites)

AUTHORS
TITLE
JOURNAL
MEDLINE
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
Structural analysis of Arabidopsis thaliana chromosome 3. II.
Sequence features of the 4,251,695 bp regions covered by 90 P1, TAC
and BAC clones
DNA Res. 7 (3), 217-221 (2000)
20363099
2 (bases 1 to 85690)
Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
Direct Submission
Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOD1
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/)
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremli1.zool.iastate.edu/cgi-bin/sp.cgi).
Genes encoding tRNAs are predicted by tRNAscan-SE
(Sean Eddy, Washington University School of Medicine, St. Louis,
http://genome.wustl.edu/eddy/tRNAscan-SE/).
This sequence may not be the entire insert of this clone. It may be
shorter because we remove overlaps between neighboring submissions.
The 5' clone is T13J10 and the 3' clone is T26G12.
Location/Qualifiers
1. .85690
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOD1"
/clone_lib="Mitsui P1"
2123. .8637
/note="contains similarity to TNP2-like transposon
protein
gene_id:MOD1.1"
/codon_start=1
/pseudo
/evidence=not_experimental
join(10671..10844,10921..11208)
/note="contains similarity to unknown protein
emb|CAA16722.1
gene_id:MOD1.2"
/codon_start=1
/pseudo
/evidence=not_experimental
15142. .16497
/note="gene_id:MOD1.3"
/codon_start=1
/evidence=not_experimental
/product="anthocyanin 5-aromatic
acyltransferase/benzoyltransferase-like protein"
/protein_id="BAB02518.1"
/db_xref="GI:11994477"
/translation="MTLHVLETARVTPDYSVINSANLKLPLTFEFLPWLFPVKR
VFVELTSTRDHFHSILPKKDSLSILNRLPLVGHITWEPEKPSIIVSNGV
VLVTIAESDADFSLSGYGORSELHALPKLPVSDSATAFSIQTLTFPNOGFSIG
RSKHAVIDGKTSSTFLKAWAQICKQLQSPENLTPSYDSLRIKTYPTLDERMIELV
VLAEDQNTSRSLTSLPSKGLDDVLATVLSRADLERLEQVKNVSPSLHSTFVI
AYAWATCFVKAGCNDRSVSLFVGDFDRDLPKLPQTVFGNCMPVGCYNKAAE
FMEKGFVTAARIELISLVKGLSSRKIETIADTFVEGFSQSWSTQFTIAGSTRGVY
EADGWRGPRVKVDIVSIQQEATAMARRDESQVEIGMCLKTKTMDSVSVSFFNGLH
S"
complement(18639..19994)

FEATURES
source

CDS

CDS

CDS

CDS

Db 15557 ATGACTCAGCGA-----CCGCTTTCTCCATCAAAATCATCGTTGTTTCCTA 15601
QY 454 GCCGAGCGCTCGGTATTGGAATAGCAAGCATCAACACCGTTAGCGACGCCCGCTGCTTC 513
Db 15602 ACCAAGGATCTCCATAGTGTGACGACACCATCGTGTATTAGACGGGAAAGCTCAA 15661
QY 514 TCGCGTTTATACGCGCTGGTCTTCAATGAGCAACA--CATTTGAAATGAAGATGAAGA 571
Db 15662 GCACCTTTTATCAAGCTTGGGCTCAAACTCTGCAACAAGAACTTCAAGACATGCCGGAGA 15721
QY 572 TGAAGAAATTAATCTTTGCGAGTTTTCGATAGATCCGTCATAAAATATCCGACCAAAATT 631
Db 15722 ATCTAACTCCGCTTTATGACCGGCTTTTGATCAAAATATCCGACCTATCTTGATGAAAGA 15781
QY 632 TGACTCCATTTATTGGA-----GAAACCGCGTAAATTTCTTTCGAATCTCGTCATCCC 686
Db 15782 TGATAGAGTAGTGAGATCTCTCAAGAGGACCAACCAACATCAGAAAGCTTGATTCGC 15841
QY 687 TC-----ATTACCGAGCGACCGCATTCGAAACACGCTTCGTTTTCACCCCAATCCA 735
Db 15842 TTCCCTCTAGTAAACTCGGAGACGATGTCGTGTAGCCACGCTAGTTCTCTCTCGGGCTG 15901
QY 736 AAATTAAGAAATTAAGAGTGGATTTCAGTCCAGAGTTTCCAAAGTTAGTCCATCTCTCAT 795
Db 15902 ATATTGAGAGACTTAGGGAACAAGTCAAGAACGTTGTCACCGAGCC---TCCACTTTGTCGA 15958
QY 796 CTTTGTAGCGATTGCAGCTTATATGCGGTGGCATAAACGAATCATTCACAGCAGATG 855
Db 15959 CTTTGTCTATTGCTTACGCTTACCGGTGAGACTGTTTCGTGAAGCAGCTGAGGATAA 16018
QY 856 AAGACCAAGACAAGAGATGATTTCTTTGATTCCGGTGCATCTAAGGCCACGATTAG 915
Db 16019 AAGACAGA-----TCAGTGAAGTCTCTGTTGCTGAGGAGATTTCAGAGACCGGTTAG 16069
QY 916 ATCCCGCGGTCCTCAAAATTTACTTCGGAAGTCTGTTATCGTACCGCTGCCGAGATGC 975
Db 16070 ATCCGAAGCTCCCGGGGACTTTACTTCGGAAGTCTGATGATTCGCGTGGGCTGTTATAAC 16129
QY 976 GCGCGGAGAGCTGGTGGGAGAGAAAGGGTGTCTTCTGGCAGCTGAGGTAATCGCGGCGG 1035
Db 16130 GTAAGCGCGGAGCTTTATGGAAGAGAGGATTTGTGACCGCGGCTGAGATTATACGG 16189
QY 1036 AGATAAAAAAAGGATCAACGACAGAGAATATTAGAAACCGGTGGAGAAATGGTCGCGG 1095
Db 16190 ATTTGGTGAAGAGGTTGAGTTCAAGAAAGATAGACACCAATTCGCGACACTTTCGTGGAG 16249
QY 1096 AGATTCGTAAGCGTTGAGAAATCATATTTT---TCGGTGCAGGATCGAGCAAGCTAG 1152
Db 16250 GATTTAGTTTTCAGAGTTGGAGTACACAGTTTGGGACATAGCCGGGTGCGACCGGTTGG 16309
QY 1153 ATCTTTACGGTGCAGATTTTGGATGGGGAAGGCGAGAGAAACGAAGAAATATTGTCGATTG 1212
Db 16310 GAGTATACAGCGGATTTCCGGTGGGGAAGACCCGTTAAAGTTGATATGTCCTCCATTG 16369
QY 1213 ATGGGAGAAATATGCAATGACGCTTTGTAAGCCAGGAGTTTTCGAAGGAGGATTGGAGG 1272
Db 16370 A---CCAAGGAGAAGCGATCCGAATGGCTGAGAGACGATGATGAGTGCAGTGGCGTTGAGA 16426
QY 1273 TTTGCTGTCTTTGCTTAAGACAAAATGGATGCTTTTTCGCTTATTTTTCACCTGGGAA 1332
Db 16427 TTGAAATGTGTTGAAAAGACTGAAATGGATTCGGTCTGCTCTCTTTTTCACCAATGGTT 16486
QY 1333 TTAATGGTTAA 1343
Db 16487 TACATAGCTAA 16497

RESULT 13
AC069470
LOCUS
DEFINITION Arabidopsis thaliana chromosome 1 clone IGF-F7F7. *** SEQUENCING IN
PROGRESS ***, 4 unordered pieces.
ACCESSION AC069470

AC069470.10 GI:12320593
HTG: HTGS-PHASE1.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (bases 1 to 136047)
AUTHORS Lin,X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Ronning,C.M.,
White,O., Nierman,W.C. and Fraser,C.M.
TITLE Arabidopsis thaliana 'IGF' BAC 'F7F7' genomic sequence near marker
'ml72'
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 136047)
AUTHORS Town,C.D. and Kaul,S.
TITLE Direct Submission
JOURNAL Submitted (01-JUN-2000) The Institute for Genomic Research, 9712
Medical Center Dr, Rockville, MD 20850, USA, cdtown@tigr.org
COMMENT On Jan 19, 2001 this sequence version replaced gi:12280790.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 4 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 100163: contig of 100163 bp in length
* 100164 100213: gap of unknown length
* 100214 112623: contig of 12410 bp in length
* 112624 112673: gap of unknown length
* 112674 116366: contig of 3963 bp in length
* 116367 116686: gap of unknown length
* 116687 136047: contig of 19361 bp in length.
FEATURES
Location/Qualifiers
1..136047
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"
/chromosome="1"
/clone="IGF-F7F7"
BASE COUNT 46028 a 22872 c 22213 g 44783 t 151 others
ORIGIN

Query Match 5.1%; Score 75.4; DB 74; Length 136047;
Best Local Similarity 46.8%; Pred. No. 1.5e-07;
Matches 623; Conservative 0; Mismatches 651; Indels 57; Gaps 10;
QY 37 ACTCGGTGGCGGAGCAATCGGTGCGCTCACATCTTCGACATGACGTGGGTGCAATTTTC 96
Db 90101 ACTCAGCCAACCTGCGACAAACTCTCTTTAACTTTCTCGACTCCATGCTACTATTCC 90160
QY 97 ATCCCATGCTTCAGTCTCTCTTACGAATTCCTTGTTCGAAGCAACATTTTTCAGAA 156
Db 90161 AGCCAGTCAAAAGAGCTCTTCTTACGAGCTCACCGAGCTACTCGTGACCACTTTCCACT 90220
QY 157 CC---ATCGTTTCAAAACTCAACATCTCTCTCTAAACACTCTATACACTTCTTCCCTC 213
Db 90221 CCAATTATCTCCCTCAAGCTCAAGGACTCTCTGCTCTATCTCGTAACTATCTCCCTC 90280
QY 214 TCTCATGCAATTTAATCTTACCTTCATCCCGGAGAAATGCCGAGGTTTCGGTATCTAT 273
Db 90281 TCATGGCCACATCATTGGGAGCCAAACGAGCCAAACCGAGCATATCGTCTCGGAA 90340
QY 274 CCGGGGACTCGGTTTCTTTTACCATCGCAGAACTTAGCGACGACTTCGATGATCTCGTCG 333
Db 90341 ACGGCGTC---GTTTGTGTGACTATAGCCGAGAGTGACGCTGACTTTTCTCTCTTTCCG 90397
QY 334 GAAATCGTCCAGATCTCCCGTTAGGCTCTACAACTTTGTCTCCCTAAATTCGCGCCCAT 393
Db 90398 GTTATGGACACGCGCTTTGTGTGACAGTTTACATGCTGTGTTCCCAAGTTACCGGTTTCTG 90457

QY	394	TCGAAGAATCCGATAGAAAACCTCTTCCAAAGTTTTCGCCGTGCAGGTGACTCTTTTCCCCAG	453
Db	90458	ATGACTCAGCGA-----CCGCTTTCCTCATACAAATCAGTGTGTTCCCTA	90502
QY	454	GCCGAGGCGTGGTATTTGAATAGCAACGATCACACGTTAGCGACGCCCGCTCGCTTTC	513
Db	90503	ACCAAGGATTTCTCATAGGTGTTGCGACACACCATCTGTTTTAGACGGGAAACGCTCA	90562
QY	514	TCGCGTTTATAACGGCTTGCTCTTCATGAGCAACA--CATTGAAATGAAGATGAAGA	571
Db	90563	GCACTTTTATCAAGCTTGGCTTCAATCTGCANACAAGAACTTCAAGATCCCGGAGA	90622
QY	572	TGAAGAATTTAAATCTTTGCCAGTTTTCGATAGATCCGTCATATAATATCCCGAGAAAT	631
Db	90623	ATCTAACTCGTCTTATGACCGGCTTTTGATCAAAATATCCGACCTATCTGATGAAAGA	90682
QY	632	TGACTCCATTTATGGA-----GAAAGGCGCTAAATTTCTTTGCTCAATCTCGTCATCC	686
Db	90683	TGATAGAGTTAGTGAGATCTCTCAAGAGAGCAACCAACATCAGAAGCTTGACTTCGC	90742
QY	687	TC-----ATTACGACGCGCGCATTCGAACACGAGTTTCGTTTTCACCCAATCCA	735
Db	90743	TTCCCTCTAGTAACTCGGAGACGATGTCGTGTAGCCAGCTAGTCTCTCTCGGGCTG	90802
QY	736	AAATTAAAGAAATTTGAAGGTTGGATTGAGTCCAGAGTTCCAAAGTTTAGTCCATCTCAT	795
Db	90803	ATATTGAGAGACTTAGGAAACAAGTCAAGAACGTCACCGACGC---TCCACTTTCGA	90859
QY	796	CTTTTGTAGCGATTGACGCTTATATGTGGCTGGCATACGAAATCATTCACAGACAGATG	855
Db	90860	CTTTTGTGATTTGCTTACGCTTACGCTGAGCTTGTTCGTGGAAGGCACTGGAGGTAATA	90919
QY	856	AAGACCAAGACAACGAGGATGCAATTTTCTTGATTCGGTTCGATTAAGGCCACGATTAG	915
Db	90920	AACACAGA-----TCAGTCAGTCTCTGTTCGTAGAGATTTTCAGAGACCGGTAG	90970
QY	916	ATCCGCGGTTCTCGAAAATTAATCTCGGAACTGCTTATGTTACGCGCTGCCGAGAATGC	975
Db	90971	ATCCGAAGCTCCCGGGGACTTACTTCGGAACTGTATGATTCCGGTGGGCTGTATAACC	91030
QY	976	GCGCGGAGAGCTGTTGGGAGCAAAAGGGGTGTTCTTCGCGAGCTGAGGTAAATCCGCGG	1035
Db	91031	GTAAGGCGCGGAGTTTATGGAAGAGAAGGGATTTGTGACGGCGCTGAGATTTATAGCG	91090
QY	1036	AGATAAAAAAAGGATCAACGACAAGAAATATTAGAAACGGTGGAGAAATGTCGCGG	1095
Db	91091	ATTTGGTGAAGGTTGAGTTTCAAGAAAGATAGAGACCATTCGCGACACTTTCGTGGAAG	91150
QY	1096	AGATTCTGAAGCGTTGCAAAAATCATATTTT---TCGGTGGCAGGATCGAGCAAGCTAG	1152
Db	91151	GATTTAGTTTTCAGAGTTGGAGTACACAGTTTGGGACGATAGCGGTCGACCCGGTTGG	91210
QY	1153	ATCTTTACGTTGCAGATTTTGGATGGGGAGCGGAGAAAGCAAGAAATATGTCGATTC	1212
Db	91211	GAGTATACGAGGCGGATTTTCGGGTGGGGAAGACCCGTTAAAGTTGATTTGTCCTCATTC	91270
QY	1213	ATGGGAGAAATATGCAATCACCGCTTGTAAAGCCAGGGATTTCCGAAGGAGGATTTGAGG	1272
Db	91271	A---CCAAGGAGAAGCGATCGCAATGGCTCAGAGACGTGATGAGTCAGGTGGCGTTGAGA	91327
QY	1273	TTTTCCTTGTCTTTCCTTAAGGACAAAATGATGCTTTTGTGCTTATTTTTCACCTGGAA	1332
Db	91328	TTGGAATGTCTTTGAAAAGACTGAAATGGATTTCGGTCGTGCTCTTTTTCACCAATGTT	91387
QY	1333	TTAATGGTTAA	1343
Db	91388	TACATAGCTAA	91398

RESULT 14
AC069470/c
LOCUS
DEFINITION Arabidopsis thaliana chromosome 1 clone IGF-F7F7, *** SEQUENCING IN

ACCESSION	AC069470	PROGRESS ***, 4 unordered pieces.
VERSION	AC069470.10 GI:12320593	
KEYWORDS	HTG; HTGS_PHASE1.	
SOURCE	thale cress.	
ORGANISM	Arabidopsis thaliana	
REFERENCE	1 (bases 1 to 136047)	Lin.X., Kaul,S., Town,C.D., Benito,M., Creasy,T.H., Ronning,C.M., Koo,H., Fujii,C.Y., Utterback,T., Barnstead,M.E., Bowman,C.L., White,O., Nierman,W.C. and Fraser,C.M.
AUTHORS	Arabidopsis thaliana 'IGF' BAC 'F7F7' genomic sequence near marker 'mi72'	
TITLE	Unpublished	
JOURNAL	2 (bases 1 to 136047)	Town,C.D. and Kaul,S.
REFERENCE	Direct Submission	
AUTHORS	Submitted (01-JUN-2000) The Institute for Genomic Research, 9712 Medical Center Dr., Rockville, MD 20850, USA, cdtown@tigr.org	
TITLE	On Jan 19, 2001 this sequence version replaced gi:12280790.	
JOURNAL	* NOTE: This is a 'working draft' sequence. It currently	
COMMENT	* consists of 4 contigs. The true order of the pieces	
	* is not known and their order in this sequence record is	
	* arbitrary. Gaps between the contigs are represented as	
	* runs of N, but the exact sizes of the gaps are unknown.	
	* This record will be updated with the finished sequence	
	* as soon as it is available and the accession number will	
	* be preserved.	
	* 1 100163: contig of 100163 bp in length	
	* 100164 100213: gap of unknown length	
	* 100214 112623: contig of 12410 bp in length	
	* 112624 112673: gap of unknown length	
	* 112674 116636: contig of 3963 bp in length	
	* 116637 116886: gap of unknown length	
	* 116887 136047: contig of 19361 bp in length.	
FEATURES	Location/Qualifiers	
Source	1..136047	
	/organism="Arabidopsis thaliana"	
	/db_xref="taxon:3702"	
	/chromosome="1"	
	/clone="IGF-F7F7"	
BASE COUNT	46028 a 22872 c 22213 g 44783 t 151 others	
ORIGIN		
Query Match	4.7%; Score 68.8; DB 74; Length 136047;	
Best Local Similarity	46.3%; Pred. No. 6.1e-06;	
Matches	461; Conservative 0; Mismatches 507; Indels 28; Gaps 6;	
QY	423 GTTTTCGCGTGCAGGTGACTCTTTTCCAGCGGAGCGTTCGATTGAATAGCAACG	482
Db	28855 GTTCTTTCTTTCAGGTCACCTTTGTTCCCAACCAAGGGTTTGTATGTGATCGCTTCT	28796
QY	483 CATCACACCCTTAGCAGCGCCGCTGTTCTCGCGTTTATAACGGCTTGGCTTCAATG	542
Db	28795 CATCATCTCTTATGACGGGAAACAGTGTAGTTTATAAAATCATCGTCCACATT	28736
QY	543 AGCAACACATTTGAAATGAAGATGAAGATGAAG-----AATTTAAATCTTT	589
Db	28735 TGTAAACATGAGCCATGGAATTTACCGGAGGATTTGACTCCGGTTTATAGATCGTACGGTT	28676
QY	590 GCAGATTTTCGATAGATCCGTCATATAAATATCCGAGAAATTTGACTCCATT-----AT	644
Db	28675 ATCATGTCCCTCAAGTCTTTGATGCAAAAATCATTTGCAACTTTTGTTCATATTTCTCAGAA	28616
QY	645 TGGAGAAACGCGCTAAATTTTCCTTTGCAATCTCGTATCCCTCATTTACCGAGCGACCG	704
Db	28615 GTGAAGATAGTTTCAGATCTNTTGAAGCTGCTTCTCCCAAGGAGATCACGCCCGACCTA	28556
QY	705 ATTGCAACACGTCGTTTTCACCCCAATCCAAAATTAAGAAATTTGAAGGTTGGATTCAG	764

Db 28555 GTACGGATCTCGCTCGAGTTGACTCGAGAGATATAGAGAACTTAGGAGCAAGCCCAAG 28496
Qy 765 TCCAGAGTTCCAAAGTTTAGTCCATCTCTCATCTTTTGTAGCGATTGACGCTTATATGCG 824
Db 28495 AGGGAGTCAGCTCGCTCATCATCATGAGTTCTACTTGTCAACGTTGTTGCCAACGG 28436
Qy 825 GCTGGCATACGAATCATTCACAGAGATGAAGACCAGACAAGCAGAGATGCATTTTTC 884
Db 28435 TATCTTTGACATGCTTGTGTAAGACGGCTGGAGAGATGAGAACAGACCGTCCGTTTC 28376
Qy 885 TTGATTCCGGTCGATCTAAGCCACGATAGATCCCGCGTTCTCGAAATTTACTTCCGG 944
Db 28375 ATGTATGCTGCTGATTTTCAGGAACCGTTAGATCCACCGTTTCCGAGATGATTTCCGG 28316
Qy 945 AACTGCTTATCG- ---TACGCGCTCCGAGAAATGCGCGGAGAGCTGGTGGGAGAGAA 1000
Db 28315 AACTCGGTGTTCCCGATAGTTGCTTTGTTACAGGCGAATGTTTTCGGAGAAAT 28256
Qy 1001 AGGGTGTTCGCGAGCTGAGGTAAATCCGCGCGGAGATAAAAAAAGGATCAACGCAA 1060
Db 28255 GGGTTTGTAAATATGTCGAAATTTCTAAGTGATTCGGTTAGAAGTATTTGGTTACGAAAG 28196
Qy 1061 GAGATATTAGAAACGGTGGAGAAATGTCGCCGAGATTCGTAAGCGTTGCACAAATC 1120
Db 28195 TTGGAACAAATTTGCGAATTTGTAATTAATGGAACAAAGACTGTGAACACAGGTACACAG 28136
Qy 1121 ATATTTTTCGGTGGCAGGATCAGCAAGCTAGATCTTTACGGTGCAGATTTTGGATGGG 1180
Db 28135 AT-TGGATCCATTGCGGTTCGAACAGTTTGGGCTATACGGGTACAGATTTCCGGTGGG 28077
Qy 1181 GAAGCGAGAAAGCAAGAAATATTTCGATGATGGGAGAAATATGCAATGACGCTTTG 1240
Db 28076 AAACCATGTAAACAGTGAGATTGCTCCATTGACCGA- ---AACGAGCGTTTTCGATGTC 28020
Qy 1241 TAAACCCAGGATTTGGAAGAGGATTGGAGTTTGTCTGCTTTCCTTAAGGACAAAT 1300
Db 28019 GGAGAGGAGGGATGAACCCGAGGTTGGAGATCGGCTTGTGTTGAAGAAATGTGAGAT 27960
Qy 1301 GGATGCTTTGCTGCTTATTTTCTACTGG- -AATTAATGGTTAATAATGATGATGAAT 1358
Db 27959 GGACATATTATTATCTGTTCAAAATGGTTTAGAAATGGTTCTTTAGGGTCTTTGATT 27900
Qy 1359 AAATAATATTATTATGTAACAAATTAATTAAGTGT 1394
Db 27899 CTAATGTTTGAAGATGAATTTTTTTTTTCTCTTT 27864

RESULT 15

166494/c 166494 7218 bp DNA PAT 28-DEC-1997
LOCUS Sequence 14 from patent US 5670367.
DEFINITION I66494
ACCESSION I66494
VERSION I66494.1 GI:2724471
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 7218)
AUTHORS Dörner, F., Schellinger, F. and Falkner, F. Gunter.
TITLE Recombinant fowlpox virus
JOURNAL Patent: US 5670367-A 14 23-SEP-1997;
FEATURES
Location/Qualifiers
1..7218
/organism="unknown"
BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others
ORIGIN
Query Match 4.4%; Score 65.8; DB 10; Length 7218;
Best Local Similarity 7.5%; Pred. No. 2.1e-05;
Matches 31; Conservative 219; Mismatches 161; Indels 0; Gaps 0;
Qy , 862 AAGACAACGAGGATGCATTTTCTTTGATTCGGTGCATCTAAGGCCACGATTAGATCCGC 921

Db 1505 AAAACGGCATCTAGCATCACTGTAAATTACCTATCTATGCAAGTAGTTAAAGAGATAGAA 1446
Qy 922 CGGTTCTCTGAAAATTACTTCGGGAACCTGTTATCTACGCGCTGCCGAGAATGCCGCGCG 981
Db 1445 GAATTTGGTACRR 1386
Qy 982 GAGAGCTGGTGGGAGAGAAAGGGGTGTTTCTGGCAGCTGAGCTAATCGCGCGGAGATAA 1041
Db 1385 RRR 1326
Qy 1042 AAAAAGGATCAACACAAGAGAAATATTAGAAACGGTGGAGAAATGGTCGCCGGGAGATTC 1101
Db 1325 RRR 1266
Qy 1102 GTAAACGCTTCAGAAATCATATTTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTAGC 1161
Db 1265 RRR 1206
Qy 1162 GTGAGATTTTGGATGGGGAAGCGAGAAAGCAAGAAATATTGCGATTGATGGGAGA 1221
Db 1205 RRR 1146
Qy 1222 AATATGCAATGACGCTTTGTAAAGCCAGGATTTTCGAAGGAGGATTTGGAGG 1272
Db 1145 RRR 1095

Search completed: November 5, 2001, 16:50:26
Job time: 10796 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:37 ; Search time 445.49 Seconds
(without alignments)
2084.596 Million cell updates/sec

Title: US-08-894-356C-4

Perfect score: 1479

Sequence: 1 CCGTGATCGAACGTGTAGA.....TTGAAAAAAAAAAAAAAAAAAAA 1479

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- N_Geneseq_0601.*
- 1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
 - 2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
 - 3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
 - 4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
 - 5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
 - 6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
 - 7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
 - 8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
 - 9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
 - 10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
 - 11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
 - 12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
 - 13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
 - 14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
 - 15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
 - 16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
 - 17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
 - 18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
 - 19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
 - 20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
 - 21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
 - 22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1478.6	100.0	1479	17 AAT37311	Aromatic acyl tran
2	760.6	51.4	1518	17 AAT37313	Aromatic acyl tran
3	216.8	14.7	1622	17 AAT37309	Aromatic acyl tran
4	190.6	12.9	1508	17 AAT37312	Aromatic acyl tran
5	175.4	11.9	1703	17 AAT37308	Aromatic acyl tran
C 6	94.8	6.4	936	22 AAF58252	Oligonucleotide D1
C 7	94.8	6.4	936	22 AAF58254	Oligonucleotide D1
C 8	94.8	6.4	936	22 AAF58257	Oligonucleotide D1
C 9	94.8	6.4	936	22 AAF58259	Oligonucleotide D2
C 10	94.8	6.4	936	22 AAF58262	Oligonucleotide D2
C 11	94.8	6.4	938	22 AAF58255	Oligonucleotide D1

12	91	6.2	936	22 AAF58252	Oligonucleotide D1
13	91	6.2	936	22 AAF58254	Oligonucleotide D1
14	91	6.2	936	22 AAF58257	Oligonucleotide D1
15	91	6.2	936	22 AAF58259	Oligonucleotide D2
16	91	6.2	936	22 AAF58262	Oligonucleotide D2
17	91	6.2	938	22 AAF58255	Oligonucleotide D1
18	76.8	5.2	1648	21 AAC64785	Lemon acyl transe
19	72.6	4.9	1379	21 AAC46852	Arabidopsis thalia
C 20	65.6	4.4	1240	21 AAC53227	Arabidopsis thalia
C 21	64	4.3	244	22 AAF58238	Oligonucleotide D1
22	62	4.2	1502	21 AAC49421	Arabidopsis thalia
23	62	4.2	1504	21 AAC35336	Arabidopsis thalia
24	61.6	4.2	244	22 AAF58238	Oligonucleotide D1
25	45	3.0	1820	20 AAZ00434	Human secreted pro
26	45	3.0	5760	6 AAN50530	Sequence encoding
27	44.4	3.0	3690	21 AAF15891	Human prostate can
28	42.8	2.9	14066	20 AAX99556	Nucleic acid seque
29	42.6	2.9	1240	13 AAQ24177	Tox2a gene. Pyemo
30	42.4	2.9	3426	22 AAC67368	Human NFAR-1 codin
31	42	2.8	580	22 AAC92918	Human ISF CDNA clo
32	42	2.8	2620	22 AAC92908	Human immune suppr
33	41.6	2.8	1556	21 AAC66427	Human secreted pro
34	41.4	2.8	1320	21 AAC42254	Arabidopsis thalia
35	40.2	2.7	1613	21 AAA46500	CDNA encoding enzy
36	40	2.7	1569	21 AAC59282	Human secreted pro
37	39.6	2.7	1082138	21 AAF22305	Arabidopsis thalia
C 38	39.4	2.7	1150	20 AAX13596	Enterococcus faeca
39	39.4	2.7	1241	19 AAV29479	Nucleic acid encod
40	39	2.6	2401	22 AAF56506	Arabidopsis CDC27A
41	39	2.6	2434	22 AAF56505	Arabidopsis CDC27A
42	38.8	2.6	2833	21 AAF15797	Human prostate can
C 43	38.8	2.6	79122	21 AAF22294	BAC containing rep
44	38.6	2.6	1217	20 AAZ24836	Human secreted pro
45	38.6	2.6	1583	21 AAF15676	Human prostate can

ALIGNMENTS

RESULT 1	
AAT37311	AAT37311 standard; cDNA to mRNA; 1479 BP.
ID	AAT37311 standard; cDNA to mRNA; 1479 BP.
XX	
XX	AAT37311;
XX	AC
XX	06-FEB-1997 (first entry)
DT	
XX	Aromatic acyl transferase coding sequence.
DE	
XX	Aromatic acyl transferase; transformation; anthocyanin pigment;
KW	plants; acylation; colour; tone; colouration; colour change;
KW	Gentiana triflora; Petunia hybrida; Perilla ocimoides;
KW	Scenecio cruentus; Lavandula angustifolia; ds.
XX	
XX	Perilla ocimoides (Clone psAT208).
OS	
XX	
XX	Key Location/Qualifiers
FT	CDS 3..1343
FT	/tag= a
FT	/product= Aromatic acyl transferase.
XX	
XX	WO9625500-A1.
XX	
XX	22-AUG-1996.
XX	
XX	16-FEB-1996; 96WO-JP00348.
XX	
XX	30-JAN-1996; 96JP-0046534.
XX	17-FEB-1995; 95JP-0067159.
XX	29-JUN-1995; 95JP-0196915.
XX	
XX	(SUNR) SUNTORY LTD.

Db 1199 tcaacatctgcggcaccaggtcgacgctcatttatgtctgttgatccgcacacgcaat 1258
 QY 1149 CTAGATCTTTAGCGTCGACATTTTGGATGGGGAGGCGAGAACGAAATATGTCG 1208
 Db 1259 gtgcatgatttgatgcagatttgggtgggaaagcttgaaagcatgaaatctgttca 1318
 QY 1209 ATTGATGGGAGAAATATGCAATGACGCTTTGTAAGCCAGGATTTTCAAGGAGGATTG 1268
 Db 1319 actaatcttcggcaacactaatatttgatctctcgtccagaagatttaaaagagcaactt 1378
 QY 1269 GAGGTTCCTGCTCTTGGCTTAAGGACAAATGATGCTTTTCTGCTTATTTTCACTG 1328
 Db 1379 gagcttggaattcttctgctaagaatagatggagcgaatttgcacacatttttaacgaat 1438
 QY 1329 GCAATTAATGCTTAATAATGATGTAATTAACATAATTAATTAATTAATTAATTA 1388
 Db 1439 ttcataaatagctcagtgaggagcccttggtaagaaaaagtggtatcaatgtataa 1498
 QY 1389 AGTGTGTGAGTAACGTGAAGA 1408
 Db 1499 aaagacagacaagttatga 1518

RESULT 4
 AAT37312
 ID AAT37312 standard; cDNA to mRNA; 1508 BP.

AC AAT37312;

XX 06-FEB-1997 (first entry)

XX Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation: anthocyanin pigment;

XX plants; acylation; colour; tone; colouration; colour change;

XX Gentiana triflora; Petunia hybrida; Perilla ocimoides;

XX Senecio cruentus; Lavandula angustifolia; ds.

XX Senecio cruentus (Clone pCAT8).

XX Key Location/Qualifiers

FT 3..1367

FT /*tag= a

FT /product= Aromatic acyl transferase.

XX W09625500-A1.

XX 22-AUG-1996.

XX 16-FEB-1996; 96WO-JP00348.

XX 30-JAN-1996; 96JP-0046534.

PR 17-FEB-1995; 95JP-0067159.

PR 29-JUN-1995; 95JP-0196915.

XX (SUNR) SUNTORY LTD.

XX Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;

PI Nakao M, Tanaka Y, Yonekura K;

XX WPI; 1996-393401/39.

DR P-PSDB; AAW04726.

XX DNA coding for aromatic acyl transferase - for transforming plants

PT which produce anthocyanin pigments and thus altering colour tone,

PT e.g. of flowers

XX Claim 4; Page 69-72; 94pp; Japanese.

XX Vectors containing DNA fragments encoding proteins of plant origin

CC with aromatic acyl transferase activity may be used to transform

CC plants which produce anthocyanin pigments. The aromatic acyl

CC transferase acylates the pigments in the flower resulting in colour

CC tone changes and allowing new colourations to be produced. Six

CC specific DNA sequences encoding aromatic acyl transferase from

CC different plants are described in AAT37308-37313.

XX

SQ Sequence 1508 BP; 442 A; 294 C; 295 G; 477 T; 0 other;

Query Match 12.9%; Score 190.6; DB 17; Length 1508;

Best Local Similarity 49.6%; Pred. No. 3.9e-40;

Matches 676; Conservative 0; Mismatches 654; Indels 33; Gaps 6;

QY 2 CGTGTATCGAAACGTGTAGAGTTGGGCGCGCGGACTCGGTGGCGAGCAATCGGTGCC 61

Db 5 catctcgaaatgccgaataatcgcccccctcgggcaccatcgccatcgctcgattc 64

QY 62 GGTACATCTTTCGACATGACGTGGCTGATTTTCATCCATGCTTCAGCTCCTCTCTTA 121

Db 65 tcttactttctcgacattcattggtactcttccctcgggtccacacatctttcttcla 124

QY 122 CGAATTCCTTCTTCCAAAGTAAATTTTCAGATCCATCGTTCCTCAAACTCAAACTC 181

Db 125 tgactttccacattctaaatcccatcttcacacatctgttcccgaggtaaaaacaa 184

QY 182 TCTCTCTAAACTCTCATACACTTCTTCTCTCTCTCATCAATTTAATCTACCTTTCATC 241

Db 185 ttatcggtcactcttcaacatttttccggttctgtagtaatttgattgatttctctaa 244

QY 242 CCCGG-----AGAAATGCCGAGTTTCGGTATCTATCCGGGGACTTCGGT 286

Db 245 cactgatggttcgggttttataaaaaaacagaaataaaacacgtttgaagggtgattcgt 304

QY 287 TTCTTTTCCATCGCAGAAATCTAGCGACACTTCGATGATCTCGTGGAAATCGTCCAGA 346

Db 305 tbtggtactttgagaatgttcttgactttaataattgacagggaaaatcactcctcg 364

QY 347 ATCTCCGTTAGGCTCTACAACTTTTGTCCCTAAATTTGCCGCCCATTTGTGCAAGAAATCCGA 406

Db 365 aaaaatgaaaaacttttaccacttgaccttcatttggaataatgcaatcaaatattatgta 424

QY 407 TAGAAATCTTCCAAAGTTTTCGCCGTGACGTGACTCTTTTCCAGCGCGAGCGCTCGG 466

Db 425 ttgcgtcacggtcccaactttttcacttcaagtgaactgtttttccgggtcgggtatac 484

QY 467 TATTGGAATAGCAACGCATCACACGTTAGCGACGCCCTCGTCTCTCGCTTTTAACTAC 526

Db 485 actaggaatgacgaatcactatgacctgtgtgacgtagcagcggttcaactttctgaa 544

QY 527 GGCTTGGTCTTCAATGAGCAACACATTTGAAATTAAGATGAAGATGAAGAAATTTAAATC 586

Db 545 aggggtgacttcgattattcaatctggttagatcggtcttttttaacgaagg---atc 601

QY 587 TTTGCCAGTTTTCGATGATGATCGTATATAAATATTCGAGCAAAATTTGACTTCCATTTATG 646

Db 602 tccaccggtttttgtagat---tgattaaacatccacacatttagatgaaaaaagggttag 658

QY 647 GAGAAACGCGCTAAATTTTCTTTCGAATCTCGTCACTCTCATACCGACGCGACCGCAT 706

Db 659 acatacaaggctcgaaagt---ttataaaccttcgagccttctgttgcctcactgataaagt 718

QY 707 TCGAACACGTTTCGTTTTCACCAATCCAAATTAAGAAATTTGAAAGGTTGGATTTCAGTC 766

Db 719 tgggtcaacggtttgttgacccgaactaataatcaatctactaaaagaaagggtcttaac 778

QY 767 CAGAGTTCCAAAGTTTACTCCATCTCTCACTCTTTTGTAGGATTTGACGCTTATATGTGGGC 826

Db 779 ccaagtgcacaaactgagtagatgctctttttagcgttaactgttggttatataatggag 838

QY 827 TGGCATAACGAATCATTC-----ACACGATGAGAACCAACAGACACGAGTGCAT 880

Db 839 ttgcatacgaaatcactcgtaaaaataggadaaaggagcgagagaggttagaaca 898

QY 881 TTTCTTGTTCGGTTCGATCTAAGGCCACGATTAGATCCGCCGTTTCTCTGAAATTTACTT 940

Db 899 gtccataatcaccattgattgctcgtcttgatccaccattccacagctactt 958
 Qy 941 CGGGAACCTGCTATGCTAGCGCTGCCGAGAAATCGCGCGGAGAGCTGGTGGGAGAA 1000
 Db 959 tggtaactcgtggtgcaccatgttcccgacccttaaaaaatgtctgttgactagcga 1018
 Qy 1001 AGGGGTGTTTCTCGCAGCTAGGTAATCGCGCGGAGATAAAAAAGGATCAACGACAA 1060
 Db 1019 tgggtatcacttggtgctaaagtaattggagagctctatatgcaaaatgatataataa 1078
 Qy 1061 GAGAAATATTAGAAACGGTGGAGAAATAGTCCCGGAGATTCTTAAAGCGTTGCCAGAAATC 1120
 Db 1079 ggacggaactctgaaagatgcgagatggcatga---acctttcatgaccccgctag 1135
 Qy 1121 ATATTTTCGTCGTCAGCATCGCAACGACTAGATCTTTACGGTCAGATTTTGGATGGG 1180
 Db 1136 gaagattggtgttgctggtacacctaagctcaacttgacacttttgatttgggtggg 1195
 Qy 1181 GAAGCGGAGAAAGCAAGAAATATTGTCGATTGATGGGAGAAATATGCAATGACGCTTTG 1240
 Db 1196 gaagccgataaagtatgagactgtttcaataga---ctataatcgtcgatttctataa 1252
 Qy 1241 TAAAGCCAGGATTCGAAGGAGGATGGAGGTTTGCTTTGCTTAAGGACAAAT 1300
 Db 1253 tgcagcaaaacatcagcacaagatcttgaattggattgagctacccgagatgcaaat 1312
 Qy 1301 GGATGCTTTTCTGCTTATTTTTCACCTGGGAATTAATGGTTAA 1343
 Db 1313 ggagggcgtttcttagcatcttttgatgaaggattagagtgcaa 1355

RESULT 5

AAT37308
 ID AAT37308 standard; cDNA to mRNA; 1703 BP.
 AC AAT37308;
 XX
 XX
 DT 06-FEB-1997 (first entry)
 XX
 XX
 DE Aromatic acyl transferase coding sequence.
 KW Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
 KW Scenecio cruentus; Lavandula angustifolia; ds.
 XX
 OS Gentiana triflora var. japonica (Clone pGAT4).
 XX
 FH Key Location/Qualifiers
 FT CDS 6..1415
 FT /*tag= a
 FT /product= Aromatic acyl transferase.
 XX
 XX WO9625500-A1.
 XX
 XX 22-AUG-1996.
 XX
 XX 16-FEB-1996; 96WO-JP00348.
 XX
 XX 30-JAN-1996; 96JP-0046534.
 PR 17-FEB-1995; 95JP-0067159.
 PR 29-JUN-1995; 95JP-0196915.
 XX
 XX (SUNR) SUNTORY LTD.
 XX
 XX Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX
 XX WPI: 1996-393401/39.
 DR P-PSDB: AAW04722.
 XX
 PT DNA coding for aromatic acyl transferase - for transforming plants

PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers
 XX
 PS Claim 4; Page 53-57; 94pp; Japanese.
 CC Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313.
 XX
 SQ Sequence 1703 BP; 512 A; 353 C; 356 G; 482 T; 0 other;

Query Match 11.9%; Score 175.4; DB 17; Length 1703;
 Best Local Similarity 49.6%; Pred. No. 3.8e-36;
 Matches 703; Conservative 0; Mismatches 666; Indels 48; Gaps 8;
 Qy 48 GAGCAATCGGTGCGCTCACATTCTTCGACATGAGTGGCTGCATTTTCATCCCATGCTT 107
 Db 78 gagttatcgcctaccggttaacattcttcgatatccccgggtgcacttgaaataagatgcag 137
 Qy 108 CAGTCTCTCTTCTACGAATTCCTTGTTCAGCAAGCAACATTTTTCAGAAATCCATCGTTCCA 167
 Db 138 tcccttctcttttaacgactttccgtaccacaacacatttcttggacactgttatccct 197
 Qy 168 AACTCAAAACAACTCTCTCTTAAACCTCTATACACTTCTTCCCTCTCTCATGCAATTTA 227
 Db 198 aatcttaagccctcttctctcactctaaacacactaccctccgcttagcggaatttg 257
 Qy 228 ATCTACCTTTCATCCCGGAGAAATGCGGAGTTTCGGTA-----TCATATCCGGGAC 281
 Db 258 tggatgcgatcaaatcggcgaatgcccgaagtcttcagtaactccccgtagagggcgac 317
 Qy 282 TCGGTTTCTTTTACCATCGCAGAAATCTAGCAGCAGCTTCGATGATCTCGTGGAAATCGT 341
 Db 318 tcgataactttgatcgttgcgagctgaccagagattttgactaccctaaaggtcatcaa 377
 Qy 342 CCAGAAATCTCCGTTAGGCTCTACAATTTGTCCTTAATTTGCGCGCCCATTTGTCGAGAA 401
 Db 378 ctggtagattccaatgatattgcagtgccctttttatgtatgccacgggttataaggacc 437
 Qy 402 T---CCATAGAAAACTCTTCCAAAGTTTTCGCGCTGAGAGTGACTCTTTTCCAGGCCGA 458
 Db 438 atgcaagactataaagtgcgtcccgctgcgtagccgtaagtaaccggttttccctaacctg 497
 Qy 459 GCGCTCGGTATTTGGAATAGCAACGCATCACACCGTTTAGCGACGCCCGCTGCTTTCTCGCG 518
 Db 498 ggcatagcgtggctctgacgggcacatcattcaattgcagatgctaaagtttggtaatg 557
 Qy 519 TTTATAACGGCTTGGTCTTCAATGAGCAAAACACATTGAAATGAAGATGAAGATGAAGAA 578
 Db 558 ttcataatgcttgggacctattaaacaaa---tttgggaagacgagcgttggttgcc 614
 Qy 579 TTTAAATCTTTGCCAGTTTTCGATAGATCCGTCATAAAATATCCGAGAAATTTGATGCC 638
 Db 615 gcgaattcttccatcttctcgatagatcataatcaagaatctgtatggcctagagagaa 674
 Qy 639 ATTTATTGGAGAAACGGCTGCTTCTTTTTCACCCCAATCCAAATTAAGAAATTTG 749
 Db 735 cctcgattccaagaaggtacgagctacatatgtcctctcccttgcgtgaaatccagaagcta 794
 Qy 750 AAG-----GGTTGGATTTCAGTCCAGGTTCCAAGTTTAGTCCATCTCTCATCTTTGTA 803
 Db 795 aagaacaaaagtactgaatctctcagaggatccgaacacgaacatacgttaacgacgttca 854
 Qy 804 GCGATTGCAGCTTATATGTGGGCTGGCATAAACGAAATCATTTACAGCAGCATGAAG----- 858

Db	855	atgacggtggaatcagta	cggaacatgcacgtgtcaaaatcaaaagatgacgtcgatcatcagag	914
Qy	859	-----ACCAAGACAAACGAGCAT	TTTTTCTTGATTCGGGTGCTGATGAAGCCA	908
Db	915	gaatcatcgaaacgacgaaaaatgagctcgag	tacttcagtttaccagcgagattgccgagga	974
Qy	909	CGATTAGATCCGCCGTTCCTCAAAAT	TACTTCGGGAACTGCTTATCGTAGCGCTCCCG	968
Db	975	ctctgacgccccgcgtcgccgactaac	cactcttgcaacgtctcgcatgctgttgca	1034
Qy	969	AGAAATCGCGCGCAGAGAGCTGTG	GGAGAGAAAGGGTGTTCCTGGCAGCTGAGGTAATC	1028
Db	1035	aasgcacacataaagagttagttggga	tataaaggctctctgttcagttgcagctatt	1094
Qy	1029	GCGCCGAGATAAAAAAAGGAT---	CAACGACAGACATAATTAGAAAAACGGTCGAGAAA	1085
Db	1095	ggagagccattgaaaagaggttgcaac	acgaaaaagcgtctctgcagatgcacaaact	1154
Qy	1086	TGTCGCCCGGAGATTGCTAAACGGT	TGCAGAAATCATATTTTCGGTGGCAGGATCGAGC	1145
Db	1155	tggttatcggaatctaaatggaatcc	cttcaaaagattcttcgggattaccggatcgcc	1214
Qy	1146	AACGTAGATCTTTACGGTCACAT	TTTCGATGGGGAAAGCGAGAAAGCAAGAAATATTG	1205
Db	1215	aagtctgattcgtagtgggtagattt	ggatgggaaagcctgcacaaatttgacattacc	1274
Qy	1206	TCGATTTGATGGGAGAAATATGCA	ATGACGTTCCTGAAGCCAGGGATTTTCGAAGGAGGA	1365
Db	1275	tcgttgattatgcagaattg---	attatgtgattcagtcacgggattttgaaaaaggt	1331
Qy	1266	TTGGAGGTTTGCTGTCTTTGCTT	GCCTAAGGACAAAATGGATGCTTTTGTGCTTATT	1325
Db	1332	gtggagattggagatacattgtc	taagatttcataatggatgcatttgcacaaactcttgaa	1391
Qy	1326	CTGGGAATTAATGTTTAATTAAT	TGATTAACCTAATATTATTATGTACAAATTA	1385
Db	1392	gaaggcttttgcctcttgca	tagctctcttaataagaaccataattctgcacaaagta	1451
Qy	1386	TTAAGTGTGAGTAACGTGAAGA	ATAATCCCTATTAT	1422
Db	1452	ccaaagctctttatgaacatac	atacacaacccctacttt	1488

RESULT	6	
AAF58252/c		
ID	AAF58252	standard; DNA; 936 BP.
XX		
AC	AAF58252;	
XX		
DT	24-APR-2001	(first entry)
XX		
DE	Oligonucleotide	D1835.
XX		
KW	Electron-transfer	group; ETM; mismatch; genotyping;
KW	gene	expression; ss.
XX		
OS	Synthetic.	
XX		
PN	WO200107665-A2.	
XX		
PD	01-FEB-2001.	
XX		
PF	26-JUL-2000;	2000WO-US20476.
XX		
PR	26-JUL-1999;	99US-0145695.
PR	17-MAR-2000;	2000US-0190259.
XX		
PA	(CLIN-)	CLINICAL MICRO SENSORS INC.
XX		
PI	Umek	RM;
XX		
DR	WPI;	2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX
PS Example 6; Page 127; 153pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETW) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;
SQ

Query Match	6.4%;	Score 94.8;	DB 22;	Length 936;
Best Local Similarity	7.1%;	284;	Match 3.5e-15;	
Matches	36;	Conservative 284;	Mismatches 186;	Indels 0; Gaps
Qy	974	GC	GC	GC
Db	850	GC	GC	GC
Qy	1034	GC	GC	GC
Db	790	GC	GC	GC
Qy	1094	GC	GC	GC
Db	730	GC	GC	GC
Qy	1154	GC	GC	GC
Db	670	GC	GC	GC
Qy	1214	GC	GC	GC
Db	610	GC	GC	GC
Qy	1274	GC	GC	GC
Db	550	GC	GC	GC
Qy	1334	GC	GC	GC
Db	490	GC	GC	GC
Qy	1394	GC	GC	GC
Db	430	GC	GC	GC
Qy	1454	GC	GC	GC
Db	370	GC	GC	GC

RESULT	7
AAF58254/C	
ID	AAF58254 standard; DNA; 936 BP.
XX	
XX	
AC	AAF58254;
XX	
XX	
DT	24-APR-2001 (first entry)
XX	
XX	
DE	Oligonucleotide D1875.
XX	
XX	
KW	Electron-transfer group; ITM; mismatch; genotyping;
KW	gene expression; ss.
XX	
XX	
OS	Synthetic.

[illegible]


```
XX 26-JUL-2000; 2000WO-US20476.
PF
XX
PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface.
XX
XX Example 6; Page 127; 159pp; English.
XX
CC The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;
SQ

Query Match 6.2%; Score 91; DB 22; Length 936;
Best Local Similarity 0.5%; Pred. No. 3.4e-14;
Matches 4; Conservative 450; Mismatches 305; Indels 0; Gaps 0;

QY 721 TTTCACCAATCCAAATTAAGAAATTAAGGGTTGGATTCCAGTCCAGATTCGAAGTT 780
DB 21 WWWWWW 80

QY 781 TAGTCATCTCTCATCTTTGTAGCGATTTCAGCTTATATGTGGCTGGCATACGAAAT 840
DB 81 WWWWWW 140

QY 841 CATTACAGCAGATGAAGACCAAGACAGAGGATGCAATTTTCTTGATTCGGTGCATC 900
DB 141 WWWWWW 200

QY 901 TAAGGCCACGATTAGATCCCGCTTCTTGAAATTAATTCGGGAAGTCTTATCGTACG 960
DB 201 WWWWWW 260

QY 961 CGCTCGCAGAAATCGCGCGGAGAGCTGGTGGGAGAGAAAGGGTGTTCCTGGCAGCTG 1020
DB 261 WWWWWW 320

QY 1021 AGCTAATCGCGCGGAGATAAAAAAGGATCAACGACAGAGAAATATTAGAAACGGTGG 1080
DB 321 WWWWWW 380

QY 1081 AGAAATGGTCGCGGAGATTCGTAAGCGCTGCAGAAATCATATTTTCGGTGCAGGAT 1140
DB 381 WWWWWW 440

QY 1141 CGACGAGCTAGATCTTTACGGTGCAGATTTTGGATGGGGAAGCGAGAAAGCAAGAA 1200
DB 441 WWWWWW 500

QY 1201 TATTGCGATGATGGGAGAAATATGCAATGACGCTTTCTTAAGCAGGAGTTTGAAG 1260
DB 501 WWWWWW 560

QY 1261 GAGGATGGAGGTTTCTCTCTTTCGCTTAAGCACAAATGGATGCTTTTCTGCTTATT 1320
DB 561 WWWWWW 620

QY 1321 TTTCAGTGAATTAATGGTTAATAAATGATGTAATTAATAACTAATATTATTATGAACA 1380
```

```
DB 621 WWWWWW 680

QY 1381 ATTAATTAAGTGTGTGAGTAACGTGAAGATAATCCCTATTATATATTATGATTGTTTC 1440
DB 681 WWWWWW 740

QY 1441 AATAAAGTGTAAAGGCTCTTTGAAAAAATAAAAAA 1479
DB 741 WWWWWW 779

RESULT 13
AAF58254
ID AAF58254 standard; DNA; 936 BP.
XX
AC AAF58254;
XX
XX 24-APR-2001 (first entry)
XX Oligonucleotide D1875.
XX
XX Electron-transfer group, ETM; mismatch; genotyping;
XX gene expression; ss.
XX
XX Synthetic.
XX
XX WO200107665-A2.
XX
XX 01-FEB-2001.
XX
XX 26-JUL-2000; 2000WO-US20476.
XX
XX 26-JUL-1999; 99US-0145695.
XX
XX 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface.
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;
SQ

Query Match 6.2%; Score 91; DB 22; Length 936;
Best Local Similarity 0.5%; Pred. No. 3.4e-14;
Matches 4; Conservative 450; Mismatches 305; Indels 0; Gaps 0;

QY 721 TTTCACCAATCCAAATTAAGAAATTAAGGGTTGGATTCCAGTCCAGATTCGAAGTT 780
DB 21 WWWWWW 80

QY 781 TAGTCATCTCTCATCTTTGTAGCGATTTCAGCTTATATGTGGCTGGCATACGAAAT 840
DB 81 WWWWWW 140

QY 841 CATTACAGCAGATGAAGACCAAGACAGAGGATGCAATTTTCTTGATTCGGTGCATC 900
DB 141 WWWWWW 200
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:04:28 ; Search time 168.74 Seconds
(without alignments)
1659.305 Million cell updates/sec

Title: US-08-894-356C-4
Perfect score: 1479
Sequence: 1 CCGTGATCGAAGCTCTAGA.....TTCAAAAAAAAAAAAAAAAA 1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCUTS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	65.8	4.4	7218	1 US-08-232-463-14	Sequence 14, Appl
2	57.2	3.9	7218	1 US-08-232-463-14	Sequence 14, Appl
3	42.6	2.9	319	1 US-07-593-657-14	Sequence 14, Appl
4	41.6	2.8	615	4 US-08-998-416-186	Sequence 186, App
5	41.6	2.8	636	4 US-08-998-416-1137	Sequence 1137, Ap
6	41.6	2.8	837	4 US-08-998-416-288	Sequence 288, App
7	39.4	2.7	1241	1 US-07-593-657-6	Sequence 6, Appli
8	39.4	2.7	1241	1 US-08-942-012B-3	Sequence 3, Appli
9	38.6	2.6	19124	2 US-08-487-826B-13	Sequence 13, Appl
10	36.4	2.5	1116	3 US-08-916-443A-16	Sequence 16, Appl
11	36.4	2.5	1422	1 US-08-319-704-5	Sequence 5, Appli
12	36.4	2.5	1991	2 US-08-415-593-40	Sequence 40, Appl
13	36.4	2.5	2128	2 US-08-415-593-39	Sequence 39, Appl
14	36	2.4	174	2 US-08-378-235B-8	Sequence 8, Appli
15	36	2.4	5852	1 US-07-867-106-2	Sequence 2, Appli
16	35.8	2.4	5852	1 US-07-867-106-2	Sequence 2, Appli
C 17	35.8	2.4	6136	1 US-08-137-252-1	Sequence 1, Appli
C 18	35.6	2.4	606	6 5202236-38	Patent No. 5202236
C 19	35.6	2.4	838	6 5202236-2	Patent No. 5202236
C 20	35.6	2.4	994	6 5202236-39	Patent No. 5202236
C 21	35.6	2.4	1236	6 5202236-4	Patent No. 5202236
C 22	35.4	2.4	4140	3 US-08-894-731-2	Sequence 2, Appli
C 23	35.2	2.4	1654	3 US-08-913-842-20	Sequence 20, Appl
C 24	35.2	2.4	1744	3 US-08-913-842-27	Sequence 27, Appl
C 25	35.2	2.4	1875	3 US-08-913-842-1	Sequence 1, Appli
C 26	35.2	2.4	1960	2 US-08-595-937A-1	Sequence 1, Appli
C 27	35.2	2.4	1965	3 US-08-913-842-2	Sequence 2, Appli

C 28 35.2 2.4 5173 1 US-08-242-677-1 Sequence 1, Appli
C 29 35.2 2.4 8920 2 US-08-446-855A-1 Sequence 1, Appli
C 30 35.2 2.4 8920 4 US-09-150-741-1 Sequence 1, Appli
31 35 2.4 1233 3 US-08-853-839-1 Sequence 1, Appli
32 35 2.4 2061 3 US-09-020-033-1 Sequence 1, Appli
C 33 34.6 2.3 837 4 US-08-998-416-288 Sequence 288, App
34 34.6 2.3 872 4 US-08-998-416-487 Sequence 487, App
C 35 34.6 2.3 2058 2 US-08-749-391-1 Sequence 1, Appli
C 36 34.6 2.3 2058 3 US-09-390-200-1 Sequence 1, Appli
C 37 34.6 2.3 3607 1 US-08-647-351B-1 Sequence 1, Appli
C 38 34.4 2.3 1680 2 US-08-700-152A-3 Sequence 3, Appli
39 34.4 2.3 2058 2 US-08-749-391-1 Sequence 1, Appli
40 34.4 2.3 2058 3 US-09-390-200-1 Sequence 1, Appli
41 34.4 2.3 4483 1 US-08-181-271A-5 Sequence 5, Appli
42 34.4 2.3 4483 1 US-08-449-315-5 Sequence 5, Appli
43 34.4 2.3 4483 1 US-08-444-803-5 Sequence 5, Appli
44 34.4 2.3 4483 1 US-08-449-043-5 Sequence 5, Appli
45 34.4 2.3 4483 1 US-08-456-265A-5 Sequence 5, Appli

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; CLONE: ptzgt-F1s
US-08-232-463-14

Query Match 4.4%; Score 65.8; DB 1; Length 7218;
Best Local Similarity 7.5%; Pred. NO. 9.7e-10;
Matches 31; Conservative 219; Mismatches 161; Indels 0; Gaps 0;
Qy 862 AAGACAGGAGGATGCAATTTCTTGATTCGGTGCATCTAAGCCACGATTAGATCGG 921
Db 1505 AAAACGGCATGAGCATCACTGTAATTAACCTATCTATGCAAGTAGTAAAGAGATAGAA 1446
Qy 922 CGGTTCTGAAATTTACTTCGGGAACGCTTATCTAGCGGCTCCGAGATCGCGCGC 981
Db 1445 GAATTTGGTACRR 1386
Qy 982 GAGAGCTGGTGGGAGAGAAAGGGTGTCTGCGAGCTGAGGTAATCGCGCGGAGATAA 1041
Db 1385 RRR 1326
Qy 1042 AAAAAGATCAACGACAGAGATATAGAACGGTGGAGAAATGGTCGCCGAGATTC 1101
Db 1325 RRR 1266
Qy 1102 GTAAAGCGTTGCAGAAATCATATTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTACG 1161
Db 1265 RRR 1206
Qy 1162 GTCAGATTTTGGATGGGAGCGGAGAGAAAGAAATATTTGTCGATTGATGGGAGA 1221
Db 1205 RRR 1146
Qy 1222 AATATGCAATGACCTTTGTAACCCAGGATTCGAAGAGGATTCGAGG 1272
Db 1145 RRR 1095

RESULT 2

US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZgpt-Fls
US-08-232-463-14
Query Match 3.9%; Score 57.2; DB 1; Length 7218;
Best Local Similarity 3.8%; Pred. NO. 4.2e-07;
Matches 14; Conservative 213; Mismatches 141; Indels 0; Gaps 0;
Qy 60 CCGCTCACATCTTCGACATGACGTGCTGCATTTTCATCCATGCTCAGCTCCTCTTC 119
Db 1089 YY 1148
Qy 120 TACGAATTCCTCTTTCACAGCAACATTTTTCAGAAATCCATGTTCCAAAACCTCAACAA 179
Db 1149 YY 1208
Qy 180 TCTCTCTCTAAACTCTCATACACTTCTTCCTCTCTCATGCAATTTAACTACCCCTCA 239
Db 1209 YY 1268
Qy 240 TCCCGGAGAGAAATCGCGAGTTCGCTATCTATCTATCCGGGACTCGGTTCTTTCACCATC 299
Db 1269 YY 1328
Qy 300 CAGAAATAGCAGACGACTCTCGATGATCTGCTCGAAATCTCCAGAAATCTCCGTTAGG 359
Db 1329 YY 1388
Qy 360 CTCTACAACTTTGTCCTAAATTCGCCGCCCATCTTCGAGAAATCCGATAGAAAACCTCTC 419
Db 1389 YY 1448
Qy 420 CAAGTTT 427
Db 1449 TATCTCT 1456
RESULT 3
US-07-593-657-14
Sequence 14, Application US/07593657
Patent No. 5268317
GENERAL INFORMATION:
APPLICANT: Tomalski, Michael D.
APPLICANT: Miller, Lois K.
TITLE OF INVENTION: INSECT-SPECIFIC PARALYTIC NEUROTOXIN
TITLE OF INVENTION: GENES FOR USE IN BIOLOGICAL INSECT CONTROL: METHODS AND
TITLE OF INVENTION: COMPOSITIONS
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: Greenlee and Winner, P.C.
STREET: 5370 Manhattan Circle, Suite 201
CITY: Boulder
STATE: CO
COUNTRY: USA
ZIP: 80303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/593,657
FILING DATE: 19901004
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Ferber, Donna M.
REGISTRATION NUMBER: 33,878

; INFORMATION FOR SEQ ID NO: 1137:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 636 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE: PAG1692RP
; ORGANISM: US-08-998-416-1137

Query Match 2.8%; Score 41.6; DB 4; Length 636;
Best Local Similarity 51.6%; Pred. No. 0.0051;
Matches 95; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
Qy 1296 AAAATGGATGCTTTTGGCTTATTTTTCACCTGGGAATTAATGTTAATAATGATGTA 1355
Db 7 AATATATAATTTAAATTTAATAAGATTAAATAAACTTTTTTATTATAATTTAAGTA 66
Qy 1356 ATTAACCTAATATTATTATGTAACAAATTAATTAAGTGTGAGTAACGTGAAGAATAATCC 1415
Db 67 TTAATATTATAACTATATATATCATTTATTATAAATTAATTTATGATTATAATAC 126
Qy 1416 CTATTATATTATGATTTGGTTCACAAATAAGTGAAGCTCTTGAAAAAATAATCC 1475
Db 127 TATTATATAATTTATATATAATTTACTTAAATTCATCATTTATTAATATTAATAT 186
Qy 1476 AAAA 1479
Db 187 AAAA 190

RESULT 6
US-08-998-416-288
; Sequence 288, Application US/08998416
; Patent No. 6239264
; GENERAL INFORMATION:
; APPLICANT: Philippsen, Peter
; APPLICANT: Pohlmann, Rainer
; APPLICANT: Steiner, Sabine
; APPLICANT: Mohr, Christine
; APPLICANT: Wendland, Jurgen
; APPLICANT: Knechtle, Philipp
; APPLICANT: Reibschung, Corinne
; TITLE OF INVENTION: GENOMIC DNA SEQUENCES OF ASHBYA GOSSYPII
; NUMBER OF SEQUENCES: 1152
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: No. 6239264artis Corporation
; STREET: 3054 Cornwallis Road
; CITY: Research Triangle Park
; STATE: No. 6239264th Carolina
; COUNTRY: USA
; ZIP: 27709
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/998,416
; FILING DATE: 24-DEC-1997
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: CH 0016/97
; FILING DATE: 31-DEC-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Meigs, J. Timothy
; REGISTRATION NUMBER: 38, 241
; REFERENCE/DOCKET NUMBER: PF/5-30306/A/CCC1976
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 919-541-8587

TELEFAX: 919-541-8689
; INFORMATION FOR SEQ ID NO: 288:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 837 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE: PAG1241RP
; ORGANISM: US-08-998-416-288

Query Match 2.8%; Score 41.6; DB 4; Length 837;
Best Local Similarity 51.6%; Pred. No. 0.0061;
Matches 95; Conservative 0; Mismatches 89; Indels 0; Gaps 0;
Qy 1296 AAAATGGATGCTTTTGGCTTATTTTTCACCTGGGAATTAATGTTAATAATGATGTA 1355
Db 7 AATATATAATTTAAATTTAATAAGATTAAATAAACTTTTTTATTATAATTTAAGTA 66
Qy 1356 ATTAACCTAATATTATTATGTAACAAATTAATTAAGTGTGAGTAACGTGAAGAATAATCC 1415
Db 67 TTAATATTATAACTATATATATCATTTATTATAAATTAATTTATGATTATAATAC 126
Qy 1416 CTATTATATTATGATTTGGTTCACAAATAAGTGAAGCTCTTGAAAAAATAATCC 1475
Db 127 TATTATATAATTTATATATAATTTACTTAAATTCATCATTTATTAATATTAATATAT 186
Qy 1476 AAAA 1479
Db 187 AAAA 190

RESULT 7
US-07-593-657-6
; Sequence 6, Application US/07593657
; Patent No. 5266317
; GENERAL INFORMATION:
; APPLICANT: Tomalski, Michael D.
; APPLICANT: Miller, Lois K.
; TITLE OF INVENTION: INSECT-SPECIFIC PARALYTIC NEUROTOXIN
; TITLE OF INVENTION: GENES FOR USE IN BIOLOGICAL INSECT CONTROL: METHODS AND
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Greenlee and Winner, P.C. 201
; STREET: 5370 Manhattan Circle, Suite 201
; CITY: Boulder
; STATE: CO
; COUNTRY: USA
; ZIP: 80303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/593,657
; FILING DATE: 19901004
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Ferber, Donna M.
; REGISTRATION NUMBER: 33, 878
; REFERENCE/DOCKET NUMBER: 14-90
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 303/499-8080
; TELEFAX: 303/499-8089
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1241 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double

```

; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; FEATURE
;
; NAME/KEY: CDS
; LOCATION: 119..985
; FEATURE:
;
; NAME/KEY: unsure
; LOCATION: -103..-100
; FEATURE:
;
; NAME/KEY: unsure
; LOCATION: 47..49
; US-07-593-657-6

```

Query Match	2.7%	Score 39.4	DB 1	Length 1241
Best Local Similarity	51.4%	Pred. No. 0.037		
Matches 91	Conservative 0	Mismatches 86	Indels 0	Gaps 0
Qy 1303	ATGCTTTTGCCTTATTTTTCAC	TGGGAATTAATGTTAATAATG	TATGTAATTAAC	1362
Db 1056	ATTTTAAATTAGTAATAATTCG	ATAATTTTAC	TATTTTAAATTTACTAAATTCG	1115
Qy 1363	TAATATTATTATGTAACAATTA	ATTAAGTTTGAGTAACGTGA	GAATAAATCCCTATTAT	1422
Db 1116	TAATTTTATTGAAGATGAC	TCTTAATAAATTTG	ATTTTTTAAACACAATAAT	1175
Qy 1423	ATATTTATGATTTGGTTCAA	ATAAGTGTAAGCCTCT	TGAAAAA	1479
Db 1176	ATATTTTATTAATTAATTAAT	TTAATTAATACAAGAT	AAAAA	1232

```

RESULT      8
US-08-942-012B-3
; Sequence 3, Application US/08942012B
; Patent No. 6235278
; GENERAL INFORMATION:
; APPLICANT: Miller, Lois K.
; APPLICANT: Lu, Albert
; APPLICANT: Dierks, Peter
; APPLICANT: Black, Bruce
; TITLE OF INVENTION: Biological Insect Control Agents Expressing
; FILE OF INVENTION: Insect-Specific Toxin Genes, Methods and Compositions
; FILE REFERENCE: 28-96a
; CURRENT APPLICATION NUMBER: US/08/942,012B
; CURRENT FILING DATE: 1997-10-01
; PRIOR APPLICATION NUMBER: 08/729,606
; PRIOR FILING DATE: 2000-10-01
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1241
; TYPE: DNA
; ORGANISM: Pyemotes tritici
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (119)..(985)
; US-08-942-012B-3

```

[illegible]

```

Db 1176 atatttttaattaaatttaattacaagataaaaaaaaaaaaaaaaa 1232

RESULT 9
US-08-487-826B-13
; Sequence 13, Application US/08487826B
; Patent No. 593827
; GENERAL INFORMATION:
; APPLICANT: Sim, Kim L.
; APPLICANT: Chitnis, Chetan
; APPLICANT: Miller, Louis H.
; APPLICANT: Peterson, David S.
; APPLICANT: Su, Xin-zhaun
; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Neda
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CPI
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-3550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
US-08-487-826B-13

```

	Query Match	2.6%	Score 38.6;	DB 2;	Length 19124;
	Best Local Similarity	53.7%;	Pred. No.	0.41;	
Matches	Conservative	0;	Mismatches	69;	Gaps 0;
Qy	1331	AATTAAATGGTTTAATAAATGCTATGTAAATTAACCTAATAATTTATTTATGTCACAATAATTAATTAAG	1390		
Dbl	15308	AAAAAAAAGGAGAAAAATTTTTTAAAAAATAATAAAAAATTTAATAAATATATAATTTTGGA	15567		
Qy	1391	TGTTGACTCAAGCGTAAGAATAATCCCTATTATATATATTTGATTTGGTTCAAATTAAGTG	1450		
Dbl	15568	TAGAAATAAAAANTGAAAGAATATCAAAAAAAAAATTAATAAAAAATTTTATATAAAAAA	15627		
Qy	1451	TAAAGCCTCTTTGAAAAAATAAAAAAAAAA	1479		
Dbl	15628	AAAAATGATTTATAAAAAANAATAAAAACAA	15656		

RESULT 10
US-08-916-443A-16
; Sequence 16, Application US/08916443A
; Patent No. 6001986

GENERAL INFORMATION:
APPLICANT: Yong Sig KIM
APPLICANT: Sun Chung PARK
APPLICANT: Soo Kyung OH
APPLICANT: Hosull LEE
APPLICANT: Jeong Woo CHO
APPLICANT: Chang H. CHUNG
TITLE OF INVENTION: Antiviral Proteins, Amaranidin 1 and 2, from
TITLE OF INVENTION: Anaranthus Viridis, DNAs Encoding Therefrom
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Kenyon & Kenyon
STREET: 1025 Connecticut Avenue, N.W., Suite 600
CITY: Washington
STATE: D.C.
COUNTRY: USA
ZIP: 20036
COMPUTER READABLE FORM:
MEDIUM TYPE: 3+ Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS 6.2
SOFTWARE: WordPerfect 6.1 Windows
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/916.443A
FILING DATE: 22 AUG 1997
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: Toffenetti, Judith L.
REGISTRATION NUMBER: 39,048
REFERENCE/DOCKET NUMBER: 1942/18
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-429-1776
TELEFAX: 202-429-0796
INFORMATION FOR SEQ ID NO: 16:
SEQUENCE CHARACTERISTICS:
LENGTH: 1116 nucleic acids
TYPE: nucleotides
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: genomic DNA
US-08-916-443A-16

Query Match 2.5%; Score 36.4; DB 3; Length 1116;
Best Local Similarity 55.6%; Pred. No. 0.29;
Matches 70; Conservative 0; Mismatches 56; Indels 0; Gaps 0;
Qy 1354 TAATTAACTAATATATTATGTACAAATTAATTAAGTGTGAGTAACGTGAAGAATAAT 1413
Db 968 TAATGAAATGCAACAATAAGTAAGATATTACTTAGGTTGTAATTTCAAAGTTGTAAC 1027
Qy 1414 CCTATTATATATTATGATTTGGTTCAAATAAAGTGTAAAGCCTCTTGAAAAAATAA 1473
Db 1028 GACTTTGTGTTTAACTACTTCTATCCCTTAAAAAATAATTGCAGGTTTCCAAAAAATA 1087
Qy 1474 AAAAAA 1479
Db 1088 AAAAAA 1093

RESULT 11
US-08-319-704-5
Sequence 5, Application US/08319704
Patent No. 5814617
GENERAL INFORMATION:
APPLICANT: Hoffman, Stephen L.
APPLICANT: Charoenvit, Yupin
APPLICANT: Hedstrom, Richard C.
APPLICANT: Doolan, Denise L.
TITLE OF INVENTION: Protective 17 kDa Malaria Hepatic and
TITLE OF INVENTION: Erythrocytic Stage Immunogen and Gene
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:

ADDRESSEE: Naval Medical R & D Command
STREET: Bldg 1, T-12, 8901 Wisconsin Avenue
CITY: Bethesda
STATE: Maryland
COUNTRY: U.S.A
ZIP: 20889-5606
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/319.704
FILING DATE: 07-OCT-1994
CLASSIFICATION: 424
ATTORNEY/AGENT INFORMATION:
NAME: A. David Spevack
REGISTRATION NUMBER: 24,743
REFERENCE/DOCKET NUMBER: 75,206
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 295-6759
TELEFAX: (301) 295-1022
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 1422 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-319-704-5
Query Match 2.5%; Score 36.4; DB 1; Length 1422;
Best Local Similarity 48.1%; Pred. No. 0.34;
Matches 103; Conservative 0; Mismatches 111; Indels 0; Gaps 0;
Qy 1265 ATTGAGGTTTGTCTGCTTTCCTAAGGACAAATGATGCTTTGCTGCTTATTTTC 1324
Db 115 ATAGCATAGATTATCTCTATATAGAAAAATATTTAAAAATATTTTATAATCTTT 174
Qy 1325 ACTGGGAATTAATGTTAATAAATGATGTAATTAACATAATATTTATGTAACAATA 1384
Db 175 TATAAATTTAATTTTATTTATTAATTTTGTATGAAAAATAATATTTATATTTCTAGAA 234
Qy 1385 ATTAAGTCTTGAGTAACCTGAGATAATCCCTATATATATATATGTTGGTTCAAT 1444
Db 235 TTTAATTTTAAATAAATAATTTAATTAATCAATAATAATAATAATAATAATAATA 294
Qy 1445 AAAGTGAAGCCTCTTGAAAAAATAATAATAATAATAATAATAATAATAATAATA 1478
Db 295 CAATATATCAATATAAGCATATAATTTAAGA 328

RESULT 12
US-08-415-593-40
Sequence 40, Application US/08415593
Patent No. 5912140
Patent No. 5912140 5776726
GENERAL INFORMATION:
APPLICANT: Whoriskey, Susan K.
APPLICANT: Quinn, Cheryl L.
APPLICANT: Tao, Nianjun
APPLICANT: Politis-Virk, Karen I.
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Pneumocystis Aminoacyl tRNA
TITLE OF INVENTION: Synthetase Genes, Tester Strains and Assays
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.

ZIP: 02173
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,593
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI94-09
TELEPHONE: 617-861-6240
TELEFAX: 617-862-9540
INFORMATION FOR SEQ ID NO: 40:
SEQUENCE CHARACTERISTICS:
LENGTH: 1991 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 29..1780
US-08-415-593-40

Query Match 2.5%; Score 36.4; DB 2; Length 1991;
Best Local Similarity 48.5%; Pred. No. 0.43;
Matches 100; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 1274 TTGCTTGTCTTGGCCCTAAGACAAAATGATGCTTTTGGCTGCTTATTTTCTACTGGGAAT 1333
Db 1780 TTGATTAATTTTAAATAAAATATTTTGAATAATATTTTCAAAATATAACTCTTATT 1839
QY 1334 TAATGCTTAATAATCTATGTAATAAATAATTAATATTAATTAATTAATTAAGTGT 1393
Db 1840 TTGCGATATATTTTGAATAAAATAATTAACAAATTTGATTTATAAATGATTAATAAT 1899
QY 1394 TGAGTACGTGAAGAATAATCCCTATTAATATTAATTAATTAATTAATTAAGTGTAA 1453
Db 1900 ACTTAAATATATTAAGTGAATTTTAAAGCTTATAAATTTCTGTAAATAATCTATG 1959
QY 1454 AGCCTCTGAAAAAATAAATAAATAA 1479
Db 1960 TATGTATTAATAAATAAATAAATAA 1985

RESULT 13
US-08-415-593-39
Sequence 39, Application US/08/415593
Patent No. 5912140
Patent No. 5912140 5776726
GENERAL INFORMATION:
APPLICANT: Whoriskey, Susan K.
APPLICANT: Quinn, Cheryl L.
APPLICANT: Tao, Nijun
APPLICANT: Politis-Virk, Karen I.
APPLICANT: Schimmel, Paul R.
TITLE OF INVENTION: Recombinant Pneumocystis Aminoacyl tRNA
TITLE OF INVENTION: Synthetase Genes, Tester Strains and Assays
NUMBER OF SEQUENCES: 46
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.
STREET: Two Militia Drive
CITY: Lexington
STATE: Massachusetts
COUNTRY: U.S.
ZIP: 02173
COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/415,593
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Brook, David E.
REGISTRATION NUMBER: 22,592
REFERENCE/DOCKET NUMBER: CPI94-09
TELEPHONE: 617-861-6240
TELEFAX: 617-862-9540
INFORMATION FOR SEQ ID NO: 39:
SEQUENCE CHARACTERISTICS:
LENGTH: 2128 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: other nucleic acid
US-08-415-593-39
Query Match 2.5%; Score 36.4; DB 2; Length 2128;
Best Local Similarity 48.5%; Pred. No. 0.45;
Matches 100; Conservative 0; Mismatches 106; Indels 0; Gaps 0;
QY 1274 TTGCTTGTCTTGGCCCTAAGACAAAATGATGCTTTTGGCTGCTTATTTTCTACTGGGAAT 1333
Db 1910 TTGATTAATTTTAAATAAAATATTTTGAATAATATTTTCAAAATATAACTCTTATT 1969
QY 1334 TAATGCTTAATAATCTATGTAATAAATAATTAATATTAATTAATTAAGTGT 1393
Db 1970 TTGCGATATATTTTGAATAAAATAATTAACAAATTTGATTTATAAATGATTAATAAT 2029
QY 1394 TGAGTACGTGAAGAATAATCCCTATTAATATTAATTAATTAATTAAGTGTAA 1453
Db 2030 ACTTAAATATATTAAGTGAATTTTAAAGCTTATAAATTTCTGTAAATAATCTATG 2089
QY 1454 AGCCTCTGAAAAAATAAATAAATAA 1479
Db 2090 TATGTATTAATAAATAAATAAATAA 2115

RESULT 14
US-08-378-235B-8/c
Sequence 8, Application US/08378235B
Patent No. 5872241
GENERAL INFORMATION:
APPLICANT: Pyle, Anna M.
APPLICANT: Michels, William J.
TITLE OF INVENTION: MULTIPLE COMPONENT RNA CATALYSTS AND
TITLE OF INVENTION: USES THEREOF
NUMBER OF SEQUENCES: 11
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper and Dunham, LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10032
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/378,235B
FILING DATE: 25-JAN-1995
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/45553
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 8:
SEQUENCE CHARACTERISTICS:
LENGTH: 174 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: RNA (genomic)
US-08-378-235B-8

Query Match 2.4%; Score 36; DB 2; Length 174;
Best Local Similarity 58.3%; Pred. No. 0.11;
Matches 63; Conservative 0; Mismatches 45; Indels 0; Gaps 0;
Qy 1343 ATAAATGATCTAATTAACCTAATATTATTATGTAACAATTAATTAAGTGTGAGTAACG 1402
Db 136 ATATGTGCCAATCAATTTATTCGTTATTATTATTAAATTAATTAATTAATTAATAATA 77
Qy 1403 TGAAGAATAATCCCTATTATATATTATGATTGTTCCAAATAAAGTG 1450
Db 76 ATATTAAACCATTAATATATATTTTATTAAATATATAATAATTG 29

RESULT 15
US-07-867-106-2
Sequence 2, Application US/07867106
Patent No. 5389526
GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
US-07-867-106-2

Query Match 2.4%; Score 36; DB 1; Length 5852;
Best Local Similarity 47.4%; Pred. No. 1.2;
Matches 108; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
Qy 1252 ATTTTCGAGGAGGATTGGAGGTTTCTGCTTTGCTTTCCCTAAGACAAAATGGATGCTTTTG 1311
Db 1814 ATTTTAAATAAATCTACTTTTTTTTTTTTTTTTTTTTTTTTAAATTTTAAAAATTTT 1873
Qy 1312 CTGCTTATTTTTTCACTGGGAATTAATGCTTAATAAATGATGTAATTAACCTAATATAT 1371
Db 1874 TTTTATTAGATCTCATAATTAATAATCAATTAATAATTAAGTATTTTAAATAT 1933
Qy 1372 TATGTAACAATAATTAAGTCTTGAGTAACGTGAAGATAATCCCTATATATATTTATG 1431
Db 1934 GCAAAAACATATAAAAACTAATAGTTTAAACCAACTTTTTTCTATTTCTTTTTTTTT 1993
Qy 1432 ATTTGGTTCAATAAAGTGAAGCCCTCTTGAAAAAATAAAAAA 1479
Db 1994 TTTTTTTTTTTACTTTGAAAAAATAAAAAAATAAAAAA 2041

Search completed: November 5, 2001, 18:04:45
Job time: 15085 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:51 ; Search time 5816.79 Seconds
(without alignments)
2403.520 Million cell updates/sec

Title: US-08-894-356C-4
Perfect score: 1479
Sequence: 1 CCGTGATCGAAACGCTGTAGA.....TTGAAAAA.....1479

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues
Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_est3:*
4: gb_est4:*
5: gb_est5:*
6: gb_est6:*
7: gb_est7:*
8: gb_est8:*
9: gb_est9:*
10: gb_est10:*
11: gb_est11:*
12: gb_est12:*
13: gb_est13:*
14: gb_est14:*
15: gb_est15:*
16: gb_est16:*
17: gb_est17:*
18: gb_est18:*
19: gb_est19:*
20: gb_est20:*
21: gb_est21:*
22: gb_est22:*
23: gb_est23:*
24: gb_est24:*
25: gb_est33:*
26: gb_est34:*
27: gb_est35:*
28: gb_est36:*
29: gb_est37:*
30: gb_est38:*
31: gb_est39:*
32: gb_est40:*
33: em_estba:*
34: em_estfun:*
35: em_esthum1:*
36: em_esthum2:*
37: em_esthum3:*
38: em_esthum4:*
39: em_esthum5:*
40: em_esthum6:*
41: em_esthum7:*
42: em_esthum8:*
43: em_esthum9:*
44: em_esthum10:*
45: em_esthum11:*
46: em_esthum12:*
47: em_esthum13:*
48: em_esthum14:*
49: em_esthum15:*
50: em_esthum16:*
51: em_esthum17:*
52: em_esthum18:*
53: em_esthum19:*
54: em_esthum20:*
55: em_esthum21:*
56: em_esthum22:*
57: em_esthum23:*
58: em_esthum24:*
59: em_esthum25:*
60: em_esthum26:*
61: em_esthum27:*
62: em_esthum28:*
63: em_estin1:*
64: em_estin2:*
65: em_estin3:*
66: em_estin4:*
67: em_estin5:*
68: em_estom1:*
69: em_estom2:*
70: em_estov1:*
71: em_estov2:*
72: em_estpl1:*
73: em_estpl2:*
74: em_estpl3:*
75: em_estpl4:*
76: em_estpl5:*
77: em_estpl6:*
78: em_estpl7:*
79: em_estpl8:*
80: em_estpl9:*
81: em_estpl10:*
82: em_estro1:*
83: em_estro2:*
84: em_estro3:*
85: em_estro4:*
86: em_estro5:*
87: em_estro6:*
88: em_estro7:*
89: em_estro8:*
90: em_estro9:*
91: em_estro10:*
92: em_estro11:*
93: em_estro12:*
94: em_estro13:*
95: em_estro14:*
96: em_estro15:*
97: em_estro16:*
98: em_estro17:*
99: em_estro18:*
100: em_estro19:*
101: em_estro20:*
102: gb_est25:*
103: gb_est26:*
104: gb_est27:*
105: gb_est28:*
106: gb_est29:*
107: gb_est30:*
108: gb_est31:*
109: gb_est32:*
110: gb_est41:*
111: gb_est42:*
112: gb_est43:*
113: gb_est44:*
114: gb_est45:*
115: gb_est46:*
116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	139.4	9.4	687	111	AW102336	sd86d06.y
2	131.8	8.9	690	155	BG599447	RG599447
3	131.4	8.9	606	113	AW221049	EST297518
4	130.8	8.8	783	155	BG584708	EST486469
5	129.6	8.8	570	142	BE922784	EST428553
6	127.8	8.6	583	113	AW221050	EST297519
7	127.8	8.6	763	155	BG591758	EST499600
8	118.2	8.0	508	173	BG097054	EST461573
9	112.4	7.6	564	142	BE921494	EST425179
10	112.4	7.6	559	155	BG598154	EST496832
11	109.8	7.4	788	164	BE187619	EST336180
12	108	7.3	574	142	BE922746	EST426515
13	106.6	7.2	482	144	BF096979	EST400516
14	106.6	7.2	591	167	BE434257	EST405135
15	104.8	7.1	487	118	AW616206	EST307245
16	104.4	7.1	575	142	BE920468	EST424237
17	94.8	6.4	518	119	AW650280	EST328734
18	94.8	6.4	563	152	BG362734	sac08a08
19	93.2	6.3	603	118	AW616119	EST296884
20	92	6.2	686	120	AW776512	EST335577
21	91.2	6.2	403	143	BF053202	EST438432
22	88.6	6.0	648	134	BF457806	BF457806
23	88.4	6.0	708	138	BE660849	7-C8 Gmax
24	87.8	5.9	497	122	AW930934	EST356777
25	85.8	5.8	307	142	BE920010	EST423779
26	85.4	5.8	368	142	BE921953	EST425722
27	85	5.7	411	142	BE922378	EST426147
28	85	5.7	414	142	BE921952	EST425721
29	82.6	5.6	654	151	BF642724	BF642724
30	82.2	5.6	656	119	AW692527	BF642724
31	82.2	5.6	677	118	AW586972	EST318595
32	80.8	5.5	597	120	AW774948	EST334099
33	79.6	5.4	630	155	BG581249	EST482982
34	79.4	5.4	386	173	BG097680	EST462199
35	79.4	5.4	422	155	BG589694	EST497536
36	77.4	5.2	456	142	BE923572	EST427341
37	76.6	5.2	490	117	AW038515	EST280198
38	76.2	5.2	694	117	AW560686	EST315734
39	76.2	5.2	829	155	BG584678	EST486439
40	75.2	5.1	409	111	AW093479	EST286659
41	72.6	4.9	404	120	AW737239	EST338666
42	72.6	4.9	543	119	AW650650	EST329104
43	72.6	4.9	560	30	AV442325	AV442325
44	69.8	4.7	403	143	BF052576	BF052576
45	69.8	4.7	649	151	BF643475	BF643475

ALIGNMENTS

RESULT	1
AW102336	
LOCUS	sd86d06.y1 Gm-c1009 Glycine max CDNA clone GENOME SYSTEMS CLONE ID: Gm-c1009-1284 5' similar to TR:Q92WB4 Q92WB4 F21M11.13 PROTEIN. i, mRNA sequence.
DEFINITION	21-NOV-2000
ACCESSION	AW102336
VERSION	GI:6072949
KEYWORDS	EST
SOURCE	soybean.
ORGANISM	Glycine max
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine. 1 (bases 1 to 687)
AUTHORS	Shoenmaker, R., Keim, P., Vodkin, L., Erpellding, J., Coryell, V., Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C.,

TITLE JOURNAL COMMENT

Wyllie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R., and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Insert Length: 2097 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 431.

FEATURES source

1. .687
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1009-1284"
/clone_lib="Gm-c1009"
/lab_host="Xl10-Gold"
/note="Vector: pBluescript II XR; Site_1: EcoRI; Site_2: XhoI; The mRNA was isolated from entire roots of 2-month-old 'Williams' plants that were greenhouse grown in 5-gallon pots. To suppress nodulation, Black Gold All-Purpose potting soil was supplemented with: 0.36g/L available phosphoric acid (P2O5), 20mg/L urea N, 0.16g/L S, 0.49mg/L B, 2.5mg/L Cu, 0.15g/L Fe, 13.53mg/L Mn, 0.26mg/L Mo, 14mg/L Zn, 20mg/L Ca, and the following nutrients in a slow-release form (Osmocote): 0.165g/L ammoniacal N, 0.185g/L nitrate N, 0.35g/L available phosphoric acid, and 0.35g/L soluble potash. No nodules were visible on the roots at harvest. Stratagene's CDNA Synthesis Kit (catalog #200401) was used to synthesize the cDNA. First-strand synthesis was performed with 5-methyl dCTP, hence the ligated cDNA is hemimethylated. Stratagene's first-strand synthesis primer was used [CAGAGAGAGAGAGAGAGACTGCTCGAG(T)-18]. After second-strand synthesis is, the cDNA ends were 'polished' with clone Pru DNA polymerase, ligated to EcoRI adapters, and phosphorylated. The XhoI site within the first-strand synthesis primer was restricted by digestion with XhoI; all XhoI sites in the cDNA would be protected by their hemimethylated status. The cDNA constructs were size-fractionated with a 400bp cutoff, using a SizeSep 400 Spin column from Pharmacia. The column eluent was then ligated into Stratagene's pBluescript II XR predigested vector (pBluescript II SK(+)) that had been digested with EcoRI and XhoI, and phosphorylated. Both the white and blue colonies appear to contain recombinant plasmids with Kaim and Dr. Virginia Coryell."

BASE COUNT 157 a 247 c 91 g 187 t
ORIGIN

Query Match 9.4%; Score 139.4; DB 111; Length 687;
Best Local Similarity 52.6%; Pred. No. 2.4e-25;
Matches 321; Conservative 0; Mismatches 286; Indels 3; Gaps 1;

QY 16 GTAGAGTTGGCGCGCGGACTCGTGGCGAGCAATCGGTGGCTCACATTTCTCG 75
DB 77 GTGAAGTTGGTCCACCACCAAGCTCACTCCCTCAACCACTTCCTCCCTCTTCG 136
QY 76 ACATGACGTGGTGCATTTTCATCCCATGCTTCAGCTCCCTCTTCTACGAATTCCTTGT 135

```

Db 137 ATATTCCAGGTTCTACTGCGACCCCATCCAAAGCATCTTCTCTATGACTTTCCCCACC 196
Qy 136 CCAAGCAACATTTTTCAGATCCATCGTTCCAAAACCTCAAAACATCTCTCTCTAAAACTC 195
Db 135 CCAAGCAACATTTTTCAGATCCATCGTTCCAAAACCTCAAAACATCTCTCTCTAAAACTC 195
Db 137 CCACACACCATTTCTCCAAACAGCAGCTTCCCATCTCTCAACACATCTTTTCCCTAAACC 256
Qy 196 TCATACACTCTTCCCTCTCTCATGCAATTTAAATTAACCTTACCTCCCGGAGAAAATGC 255
Db 257 TCCAAACACTTCTCCCTCTCTCCAAATCTCATGTTCTCTCCACAGCAACCCCATCTCT 316
Qy 256 CGGAGTTTCGTAATCTATCCGGGAGCTCGGTTCTTTTCCACCATCCGAGAACTAGCGAGC 315
Db 317 CTCACATAGCTTACCTTGATGGAGACTCTCTCTCTCCACGTTGCAGAGTCCACCGCAG 376
Qy 316 ACTTCGATGATCTCTCGGAAATCGTCCAGAAATCTCCCGTTAGGCTCTACAACTTTTGCC 375
Db 377 ACTTCACCTCTCTACATCAGATTACCCACAGAGAGTTCCAAATGGCACCCTCTTGTTTC 436
Qy 376 CTAAATTCGCGCCCATTTCTGAAGAAATCCGATAGAAAAC----TCTTCCAAAGTTTTCGCG 432
Db 437 CCGCGTTTCTTACCACCGCTGTGACCAAGATGGCGCACGTGTCTCTCTTATGGCCA 496
Qy 433 TGCAGGTGACCTTTTCCAGGCGGAGCGTGGTATTTGGAATAGCAACGATCACACCG 492
Db 497 TTCAGTTACAAATTTTCCCAAGTCTGGCTTCACCATATGCTCACCTTCAACCACTTG 556
Qy 493 TTAGCGAGCGCGCGTCTCTCGGTTTATAAGCGCTTGGTCTTCAATGAGCAACACA 552
Db 557 CCAGGAGCGCAAGTCACTTCAACATTCATCAAGTTCTGGGCTCTCTTTGCANAGCAN 616
Qy 553 TTGAAATGAAGATGAAGATGAAGAAATTTAAATCTTTGCCAGTTTTCGATAGATCCGTC 612
Db 617 AAGGAAACATGGCTTNNCTTCAACCTCTGCTCTCTCTTCTCTCATGAGAGGGACAAAG 676
Qy 613 TAAATATCC 622
Db 677 TTAAGATCC 686

RESULT 2
BG599447
LOCUS BG599447 690 bp mRNA EST 12-APR-2001
DEFINITION EST504342 cSTS Solanum tuberosum cDNA clone cSTS25122 5' sequence, mRNA sequence.
ACCESSION BG599447
VERSION BG599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 690)
van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chleminogo,A.,
Bougri,O., Bueli,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
FEATURES
Location/Qualifiers
1..690
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS25122"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:

```

```

XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN
Query Match 8.9%; Score 131.8; DB 155; Length 690;
Best Local Similarity 52.5%; Pred. No. 2.2e-23;
Matches 289; Conservative 0; Mismatches 262; Indels 0; Gaps 0;
Qy 15 TGTAGAGTTGGCGCCCGGAGCTCGGTGGCGGAGCAATCGGTCCCGCTCACATTCCTC 74
Db 6 TGTCAAGTTGGCGCCACCTCCGCGGCGCGAGGAGTGACACTCCCTCTTACTTATTTT 65
Qy 75 GACATGAGTGGGTGCAATTTTCATCCCATGCTTACGCTCTCTCTAGCAATTCCTTGT 134
Db 66 GATCATGTTTGGTTAGGTTCCACCGTATAGCGGAGATATTATTCTACAAGCTCCCAAT 125
Qy 135 TCCAAGCAACATTTTTCAGAAATCCATCGTCCAAAACCTCAAAACATCTCTCTAAAAC 194
Db 126 TCCAAACCCGATTTGTTCAAAACATATTCTCTCTCTTAAAATTCATCTCTCCCTCACT 185
Qy 195 CTCATACACTCTTCTCCCTCTCTCATGCAATTTAATCTACCCCTTCATCCCGGAGAAAATG 254
Db 186 CTCAAACACTATACGCCCTTAGCTGGAACGTTGCTCTCCACTAGATACAACGGATAT 245
Qy 255 CCGAGATTTCCGTTATCTATCCGGGAGCTCGGTTCTTTTACCATCGCAGAACTCTAGCAG 314
Db 246 CTTGAGTTACATATTGTGACAGGAGATCTCTATCTGTTACTTTTCTGAGACTGATATG 305
Qy 315 GACTTCGATGATCTCGTCGGAAATCGTCCAGAAATCTCCGTTAGGCTCTACAACCTTGT 374
Db 306 GATTTCAATTTATCTCATTTGGTGACCATCCGCGTAATGCTAAGGATTTTATACACTTTGT 365
Qy 375 CCTAAATTCGCCGCCATTTGTCGAAGAATCCGATAGAAAACCTTCCAAAGTTTTCGCCGTG 434
Db 366 CTAAGTTAGGGAACCTTAAGATGACCCCGGGTCCAAATAGCTCCCGGTCTTAGCCATT 425
Qy 435 CAGGTGACTCTTTTCCAGCGCGGAGCGTCTGGTATTTGGAATAGCAACGATCACACCGTT 494
Db 426 CAAGTGACACTTTTTCGGAATCTTTGGTGTATCCATTTGTTTCTCACTAACCACTATTTGT 485
Qy 495 AGCAGACCCCGTCTGTTTCTCGCTTTTATAACGGCTTTCATGAGCAACACATTT 554
Db 486 GGTATGGAGCTACCATAGTAGGTTTCATTAGGGCGTGGGCTCTACTCCATAAATTCGGT 545
Qy 555 GAAATGAAGA 565
Db 546 GGAGATGAACA 556

RESULT 3
AW221049
LOCUS AW221049 606 bp mRNA EST 07-DEC-1999
DEFINITION EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF3F1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
1 (bases 1 to 606)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ann,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
TITLE

```

JOURNAL COMMENT

Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel.: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES source

Location/Qualifiers
1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F1"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the lcm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"
147 a 143 c 123 g 193 t

BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN
Query Match 8.9%; Score 131.4; DB 113; Length 606;
Best Local Similarity 52.0%; Pred. No. 2.8e-23;
Matches 294; Conservative 0; Mismatches 271; Indels 0; Gaps 0;

QY 1 CCGTGATCGAAGCTGTAGAGTTGGCCGCCCGGACGTCGGTGGCGGAGCAATCGGTGC 60
DB 11 CCGTGATGAGCAATGTCAGTTGGCCGCCCGGACGTCGGTGGCGGAGCAATCGGTGC 70
QY 61 CGTCTACATCTTCGACATGACGCTGCTGCTATTCATCCATCGTTCAGCTCTCTCTCT 120
DB 71 CTTCTACTATTTTGACCATGTTGGTTAGGTTTCGCCGTATGAGCGGATATATTTT 130
QY 121 ACGAATTCCTGTTCCAGCAACATTTTTCAGAAATCCATCGTTCAGAAATCCAGCAAT 180
DB 131 ACAAGCTCTCCATTTTCAAAACCGATTTGCTTCAAAACATTTATTCCTCTTAAAAAT 190
QY 181 CTCTCTTAAACTCTCATACACTTCTCCCTCTCTCATGCAATTTAATCTACCTTCAT 240
DB 191 CACTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCCGGAACGTTGCTTGTCCACTAG 250
QY 241 CCGCGAGAAAATGCGGGAGTTTCGGTATCTATCCGGGAGCTCGGTTCTTCTTCCACCATCG 300
DB 251 ATACAAACGATATCTGAGTTAGCTATGTATGTACAGAGATTTCTGTCTGTACTTTT 310
QY 301 CAGAAATCTAGCGAGCACTTCGATGATCTCGTCCGGAATCGTCCAGAAATCTCCGTTAGGC 360
DB 311 TCAGAGCTGATATGAATTTCAATATATCTCATTTGTCGACCATCCGCGTAAAGGATT 370
QY 361 TCTACAACTTTGCTTAAATTCGCCCCCATTTGTCGAGAAATCCGATAGAAACTCTTC 420
DB 371 TTTATCACTTTGTTCTTAAGTTAGGGGAACCTAAGGATGACCGGGGTCCAACATGCC 430
QY 421 AAGTTTTTCGCGGTGACGTTCTTTTCCAGCGCGGAGCGTTCGGTATTTGGAATAGCAA 480
DB 431 CGCTCTTAGCCATTCAGGTGACACTTTTCCGAATCTTGGTGTATCCATTTGGTTTCACTA 490
QY 481 CGCATCACACCGTTAGCGAGCGCCCGTTCGCGTTTATACGCGTTTGGTCTTCAA 540
DB 491 ACCATCATGTTGTTGGTATGGAGCTACTATAGCAGGGTTTCATTAAGCGGTGGGCTCTAC 550
QY 541 TGAGCAACACATTTGAATGAAGA 565
DB 551 TCCACAAATTCGGTGGACATGAACA 575

RESULT 4

LOCUS BG584708
DEFINITION EST486469 MHAM Medicago truncatula/Glomus versiforme mixed EST library cDNA clone PMHAM-19K14 5' end, mRNA sequence.

ACCESSION BG584708
VERSION 1
KEYWORDS
SOURCE Medicago truncatula/Glomus versiforme mixed EST library.
ORGANISM Eukaryota; mixed EST libraries.
REFERENCE 1 (bases 1 to 783)
AUTHORS Harrison,M.J., Liu,J., Town,C.D., Van Aken,S., Utterback,T., Cho,J.

TITLE ESTs from roots of Medicago truncatula after colonization with Glomus versiforme, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Harrison M.J.
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73401
Tel: 580-223-5810
Fax: 580-221-7380
Email: mjharrison@noble.org
Noble EST name: N380179e TIGR sequence name: MTDBT67TK More information is available at: http://www.medicago.org
Seq primer: SKmod (CTA GAA CTA 9tg GAT CC).

FEATURES source

Location/Qualifiers
1..783
/organism="Medicago truncatula/Glomus versiforme mixed EST library"
/cultivar="Medicago truncatula genotype A17"
/db_xref="taxon:119092"
/clone="PMHAM-19K14"
/clone_lib="MHAM"
/tissue_type="roots colonized with Glomus versiforme"
/dev_stage="Roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glomus versiforme. The library was made from a mixture of RNA from each of these stages."
/lab_host="E. coli strain XL0LR"
/note="Vector: pBlueScript SK-; Site_1: EcoRI; Site_2: XhoI; cDNA was prepared from polyA+ enriched RNA from roots harvested at 10, 17, 22, 31 and 38 days post-inoculation with Glomus versiforme. The cDNA was directionally ligated into the Unizap XR vector from Stratagene and packaged using Gigapack III Gold packaging extracts. Plasmids containing cDNA inserts were excised from the recombinant lambda-Zap phage using Ex-assist helper phage and propagated in XL0LR cells."
231 a 219 c 97 g 236 t

BASE COUNT ORIGIN

Query Match 8.8%; Score 130.8; DB 155; Length 783;
Best Local Similarity 53.3%; Pred. No. 4.2e-23;
Matches 323; Conservative 0; Mismatches 277; Indels 6; Gaps 2;
QY 3 GTGATGAAACGTTAGAGTTGGCCGCCCGGACGTCGGTGGCGGAGCAATC---CGTG 59
DB 9 GTGATGAAACATAGTCAAGTTGGCCACCACCAAAATTCACCTACCAACCAATTTCTA 68
QY 60 CCGCTCACATCTTCGACATGACGTGGCTGCATTTTTCATCCCATGCTTCAGCTCCTCTTC 119
DB 69 CCCTTAATCTTCTCGACATACCATGGTTTATTTGTCCCAATTCACCGTATCTCTTC 128
QY 120 TAGCAATTCCTTGTTCGAAACCAATTTTTCAGAAATCCATCGTTCGAAACCAACAA 179
DB 129 TACCAATTTCTCATCAACTCATCATTTCTTCTCAACAAACACTTCCCATCTCTCAACAC 188
QY 180 TCCTCTCTTAAACTCTCATACACTTCTCCCTCTCTCATGCAATTTAATCTACCTTCA 239
DB 189 TCACCTTCCATAACCCCTCCCTACACTTCTCCCTTTTCTTCCCTTAACCTCATCGTCCCTCA 248

```

QY 240 ---TCCCGGAGAAAATGCCGGAGTTTCGGGTATCTATCCGGGACTCGGTTTCTTTTACC 296
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 249 AATTCCCAAAACACCGATCCTTACATAGCTTACCTCGATGAAGACTCTATCTCTTACC 308
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 297 ATCGCAGATCTAGCGAGCACTTCGATGATCTCGTCGGAATCGTCCGAATCTCCGGT 356
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 309 GTAGTGAATCCTCAGCAGACTTCAATATCTCGTATCTGATTCACAAGATGCTCAAAA 368
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 357 AGGCTCTCAAACTTTGTCCTTAAATTCGCCGCCATTTGCGAAGATCCGATAGAAAATC 416
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 369 TGGCACCATCTTGTTCRAATTTACTCTCAGCTCGTACCGAACAAGACAACTCGCATC 428
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 417 TTCCAAGTTTTCGGCGTGCAGTGACTCTTTTCCAGGCCGAGCGTGGTATTGGAATA 476
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 429 ATCCCTATAATGGCTATTCAAGTCAACCGTCTTCCAAACTCCGGCTGTCTATTTCGCTA 488
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 477 GCAACGCATCACACGTTAGGACGCCCGGCTGTTTTCGGGTTTATACGGCTTGGTCT 536
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 489 AGTTACAACCATGTTGCTGCGACAGCGAAATCACTTCCACATTTTCGTAATTTTGGGCC 548
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 537 TCAATGAGCAAAACATTTGAAATCAAGATGAAGATTTTAAATCTTTGCGCAGTT 596
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 549 TCTCTTTCAAAAGCCATAGCAAAAACAAATAATTTGCTCTTGAATCTCTCCGCCT 608
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 597 TTCGAT 602
      |||||
Db 609 CTTGAT 614

RESULT 5
BE922784
LOCUS BE922784 570 bp mRNA EST 02-OCT-2000
DEFINITION EST426553 potato leaves and petioles Solanum tuberosum cDNA clone
          C5B22E2 5' sequence, mRNA sequence.
ACCESSION BE922784
VERSION BE922784.1 GI:10448860
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
          van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,
          Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning
          , C.M., Fry, W.E., Tanksley, S.D. and Baker, B.
          Generation of ESTs from potato leaves and petioles
          Unpublished (2000)
          Contact: Cathy Ronning
          The Institute for Genomic Research
          For clone request: please contact Research Genetics, Libraries
          Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
          Location/Qualifiers
          1..570
             /organism="Solanum tuberosum"
             /cultivar="Kennebec"
             /db_xref="taxon:4113"
             /clone="C5B22E2"
             /clone_lib="potato leaves and petioles"
             /tissue_type="leaflets and petioles"
             /dev_stage="8 weeks old plants"
             /lab_host="SOLR"
             /note="vector: pBlueScript SK(-); Site 1: EcoRI; Site 2:
             XhoI. Tissue was supplied by Dr. Fry (Cornell University).
             Leaflets and petioles were isolated from 8 week old
             greenhouse grown plants. The plants were watered and
             fertilized freely. The tissue was immediately frozen in
             liquid nitrogen."
BASE COUNT 139 a 136 c 115 g 180 t
ORIGIN

```

```

Query Match 8.8%; Score 129.6; DB 142; Length 570;
Best Local Similarity 52.6%; Pred. No. 8e-23;
Matches 282; Conservative 0; Mismatches 254; Indels 0; Gaps 0;

QY 1 CCGTCATCAAAACGTGTAGAGTTGGCGCCGCCGCGGACTCGTGGCGGAGCAATCGGTGC 60
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 31 CTTGATAGACAATGTCAAGTTGCGCACTCCCGCGGCGCAGCGAGGTGACACTCC 90
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 61 CGCTCACATCTTCCACATGACGTGGCTGCAATTTTCATCCCATGCTTCAGTCTCCTCTCT 120
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 91 CTCCTACTATTTTTCATGATGATGTTGGTTAGGTTCCACGGTATAGCGGATATATTCT 150
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 121 ACGAATTCCTTGTTCGAAGCAACATTTTTCAGAAATCCATCTTCCAAACATCAAAACAT 180
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 151 ACAAGCTCCCATTTTCCAAACCCGATTTTCGTTCAAAACATTTATCTCTCTTAAAAAT 210
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 181 CTCCTCTAAACACTCTCATACACTTCTTCCCTCTCTCATGCAATTTAATCTACCTTCT 240
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 211 CACTCTCCCTCACTCTCAAAACACTATAGCCCTTACGTGGAAACGTTGCTCGTCCACTAG 270
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 241 CCCGGAGAAAATGCCGAGTTTCGGTATCTATCCGGGACTCGGTTTCTTCCACCATCG 300
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 271 ATCAAAACGGATATCCTGAGTTACATTTATGTGACAGGAGATTTCTGTATCTTATCTTT 330
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 301 CAGAACTCTAGCGAGCACTTCGATGATCTCGTCGGAATTCGTCAGAAATCTCCCGTTAGCG 360
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 331 CTGAGACTGATATGATGATTTCAATTTATCTATTTGTTGACCTCCGCGTAACTAAGGATT 390
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 361 TCTACAACCTTTGTCCTTAAATTCGCGCCCATTTGTCGAGAATCCGATAGAAACATCTTCC 420
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 391 TTTATCACTTTGTTCTCTAGTTAGGGGAACCTAAGGATGCAACCGGGTCCAAATAGCTC 450
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 421 AAGTTTTCGCGTGCAGTGACTCTTTTCCAGGCGGAGGCTCGGTATTTGGAATAGCAA 480
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 451 CGGCTTAGCACTTCAAGTGACACTTTTTCGAATCTTGGTGATCCATGGTTTCACTA 510
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY 481 CGCATCACACCGTTAGCGACGCCCGCTGTTTCTCGCGTTTATTAACGGCTTGGTCT 536
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db 511 ACCATCATGTTGTTGGTGATGGAGCTACCATAGTAGGTTTCATTAGGGCGTGGGCT 566
      |||||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

RESULT 6
AW221050
LOCUS AW221050 583 bp mRNA EST 07-DEC-1999
DEFINITION EST297519 tomato fruit mature green, TAMU Lycopersicon esculentum
          cDNA clone cLEF3f3, mRNA sequence.
ACCESSION AW221050
VERSION AW221050.1 GI:6532734
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
          Lycopersicon.
          1 (bases 1 to 583)
          Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,
          Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M.
          , Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
          Generation of ESTs from tomato fruit tissue
          Unpublished (1999)
          Contact: David Frisch
          Clemson University Genomics Institute
          Clemson University
          100 Jordan Hall, Clemson, SC 29634, USA
          Tel: 864 656 4366
          Fax: 864 656 4293
          Email: dfri@clcmson.edu
          5 prime sequence.
          Location/Qualifiers
          1..583
             /organism="Lycopersicon esculentum"
             /cultivar="TA496"

```

/db_xref="taxon:4081"
/clone="cLEF3F3"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the lcm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp."
BASE COUNT 142 a 140 c 117 g 184 t
ORIGIN

Query Match 8.6%; Score 127.8; DB 113; Length 583;
Best Local Similarity 51.7%; Pred. No. 2.3e-22;
Matches 291; Conservative 0; Mismatches 272; Indels 0; Gaps 0;
QY 1 CGGTGATCGAAGCTGTAGAGTTGGCGCGCGGACTCGGTGGCGGAGCAATCGGTGC 60
DB 11 CGGTGATGAGCAATGTCAGTTGGCCCACTCCCGCGCGCAACGGAGGTGATCTCC 70
QY 61 CGGTCAATCTTTCGACATGACGTGGCTGCATTTTCATCCCATGCTTCAGCTCCTCTTCT 120
DB 71 CTCTTACTTATTGACCATGTTGCTTAGGGTTTCGCGGTATGAGCGGATATTATT 130
QY 121 AGAATTCCTTGTTCAGCAACATATTTTCAGAAATCCATCGTTCCAAACATCAACAAT 180
DB 131 ACAAGCTCTCCATTTTCAACACCGAATTCGTTCAAAACATATTTCCTCTTAAAAAT 190
QY 181 CTCTCTCTAAACTCTCATACACTCTTCCCTCTCTCATGCAATTTAATCTACCCCTCAT 240
DB 191 CACTCTCCCTCACTCTCAACACATATACGCCCTTAGCCGGAACGTTGCTGTCCACTAG 250
QY 241 CCCCAGAGAAATGCGGAGTTTCGGTATCTATTCGCGGAGCTCGGTTCTTCCACCATCG 300
DB 251 ATACAACGGATATCTCGATTAGTTATGTGACAGAGATTTCTGTCTGTACTTTT 310
QY 301 CAGAACTAGCGAGCACTTCGATGATCTCGTCGGAATCGTCAGAAATCTCCGTTAGGC 360
DB 311 TCGAGACTGATATGAATTTCAATATTCTATTTGTTGACCACTCCCGCTAAGGAT 370
QY 361 TCTACAACTTTTGCCTAAATTCGCCGCCATTCGTCGAAGAAATCCGATAGAAACTCTTC 420
DB 371 TTTATCACATTTGTTCTCAAGTTAGGGGAACCTAAGGATGACCGGGGTCCAACTAGCCC 430
QY 421 AAGTTTTCCCGTGCAGTGCACTCTTTTCCAGCGGAGCGTCGGTATTGGAATAGCAA 480
DB 431 CGCTCTTAGCCATTAGGTGACACTTTTTCGGAAATCTTGGTATCCATTGGTTTCACTA 490
QY 481 CGATCACACCGTTAGCGAGCGCCCGTCTGTTCTCGCGTTTATTAACGGCTTGGTCTCAA 540
DB 491 ACCATCATGTTGTTGTGTGGAACCTATATAGCATGGTTTCAATTAAGCGGTGGGCTCTAC 550
QY 541 TGAGCAACACATTTGAATGAA 563
DB 551 TCCACAAATACGTGGACATGAA 573

RESULT 7
BG591758 763 bp mRNA EST 12-APR-2001
LOCUS EST499600 P. infestans-challenged leaf Solanum tuberosum cDNA clone
DEFINITION BPL10A14 5' sequence, mRNA sequence.
ACCESSION BG591758
VERSION BG591758.1 GI:13609898
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

FEATURES
Location/Qualifiers
Source

Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 763)
Zhang, P., Hernandez, M., Tornqvist, C.-E., Wirtz, U., Loukolanov, A., Rangel, P., Haberland, G. T., Cho, J., Chiemingo, A., Bougri, O., Buell, C. R., Roming, C. M., Helgeson, J. and Baker, B.
Generation of ESTs from Potato Leaves Challenged with Phytophthora infestans, Incompatible Reaction
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
1..763
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="BPL10A14"
/clone_lib="P. infestans-challenged leaf"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI; whole plants were challenged with 450,000 sporangia/ml P. infestans US-1(US 940501) in Biotron (Madison, Wisconsin). Leaf tissue was collected at 1, 2, 5, 12, and 24 hours post-challenge and frozen in liquid nitrogen immediately upon removal. Kennebec plants showed no signs of HR. Katabdin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

BASE COUNT 212 a 162 c 156 g 233 t
ORIGIN

Query Match 8.6%; Score 127.8; DB 155; Length 763;
Best Local Similarity 53.2%; Pred. No. 2.5e-22;
Matches 293; Conservative 0; Mismatches 257; Indels 1; Gaps 1;
QY 16 GTAGAGTTGGCGCGCGCGGAC-TCGGTGGCGGAGCAATCGGTGGCGCTCACATTCTTC 74
DB 2 GTCAAGTTGGCGCACCTCCGCGCGCGGAGGTGACACTCCCTCTACTTATTTT 61
QY 75 GACATGACGTGGTGCATTTTCATCCCATGCTTCAGTCTCTCTTACGAATTCCTTGT 134
DB 62 GATCATGTTTGGTTAGGTTTCACCGGTATACGGCGGATATTATTCTACAAGCTCCCAT 121
QY 135 TCCAAGCAACATTTTTCAGAAATCCATCGTTCCAAACTCAACAACTCTCTCTAAACT 194
DB 122 TCCAAACCCGATTTTCGTTTCAAAACATATTTCCTACTTTAAAAATTCACCTCCCTCA 181
QY 195 CTTCATACATCTTCTCCTCTCTCATGCAATTTAATCTACCTTCATCCCGGAGAAATG 254
DB 182 CTCAACACTACACGCCCTTAGCGGAAATGTTGCTGCTCCACTAGATACAAATGGATAT 241
QY 255 CCGGAGTTTCGGTATCTATCCGGGAGCTCGGTTTCTTCCATCGCAGAACTTAGCGAC 314
DB 242 CCTGAGTTAGGTTATGTGACGGAATTTCTGTATCTGTACTTTTGTGAGACTGATATG 301
QY 315 GACTTCGATGATCTCGCGGAATCGTCCAGAAATCTCCCGTTAGGCTTACAACTTTGTC 374
DB 302 GATTTCGAATTTATCTATTTGGTACCCTCCGCGTAATGCTAAGGATTTTATCATCTTGT 361
QY 375 CCTAAATTCGCCCATTTGTCGAAGAAATCCGATAGAAAATCTCTCCAAAGTTTCGCCG 434
DB 362 CCTAAGTTAGGGGAACCTAAGGATGACCGGGGTCCAATAGCTCGGCTTAGCCATT 421
QY 435 CAGGTGACTCTTTTCCAGGCCGAGCGTGGTATTGGAATAGCAACGATACACCGTT 494
DB 422 CAAGTGACACTTTTTCGGAATCTTGGTGTATCCATTGGTTTCTACTAACCATCATGTTGT 481

Db 166 TTGGTTCAAAACCAATATTCTTACTTAAACATTTCACTCTCCCTCACTCTCAAAACACTAC 225
 Qy 207 TTCCCTCTCTCATGCAATTAATCTACCTTTCATCCCGGAGAAATGCGGAGTTTCGG 266
 Db 226 ACACCTTTAGCTGGCAACATGCTTGTCCACATAAATTCGAGTGGTTATCCTGAGTTGCGT 285
 Qy 267 TATCTATCCGGGAGCTCGGTTCTTTTCCACATCGCAGAAATCTAGCGAGCTTTCGATGAT 326
 Db 286 TATGTGACTGGAGATCTATATCTTGTACTTTTATTTGAGACTGATATGAATTTCAATCAT 345
 Qy 327 CTGCTGGGAAATCGTCCGAAATCTCCGTTAGGCTCTACAATTTGTCCTTAATTTGCGG 386
 Db 346 CTCAATTTGGTAACCATCTCTGAAAT---GCTAAGGATTTTATCCCATCTCTCAATTTGGCA 402
 Qy 387 CCATTTGTGCAAGAAATCCCATGAAACACTTCCCAAGTTTTCGCGTGCAGTGACTCTT 446
 Db 403 CAACCTAAGATGACACGGGGTCAAACTAGTCCCGGTCTTACCCATTTCAAGTGACACTT 462
 Qy 447 TTCCAGGCGGAGCGCTCGGTATTGGAATAGCAACGCATCACACCGTTAGCGACGCCCG 506
 Db 463 TTCCGAATCTTGGCATATCCGTTGGTTTATAGTAACCATCATCTCGCTTGTGATGGAAT 522
 Qy 507 TCGTTTCTCGGTTTATACGCGTTGGCTTTCATGAGCAAA 548
 Db 523 ACCATCGTGAATTTCAAGAACATGGGTTTACTCAACAAA 564

RESULT 10
 BG598154 659 bp mRNA EST 12-APR-2001
 LOCUS EST496832 cSTS Solanum tuberosum cDNA clone cSTS20M23 5' sequence,
 mRNA sequence.
 ACCESSION BG598154
 VERSION BG598154.1 GI:13616294
 KEYWORDS EST.
 SOURCE potato.
 ORGANISM Solanum tuberosum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
 REFERENCE 1 (bases 1 to 659)
 Bouari O., Buell C.R., Ronning C., Tanksley S. and Baker B.,
 Generations of ESTs from sprouting potato eyes
 Unpublished (2000)
 Contact: Cathy Ronning
 The Institute for Genomic Research
 For clone info: please contact Research Genetics, Libraries
 Division tel 1-800-711-6195, email cdna@resgen.com
 Seq primer: M13F-R.
 FEATURES
 Location/Qualifiers
 1..659
 /organism="Solanum tuberosum"
 /cultivar="Kennebec"
 /db_xref="taxon:4113"
 /clone="cSTS20M23"
 /clone_lib="cSTS"
 /tissue_type="sprouting eyes from tubers"
 /dev_stage="12-14 weeks post harvest"
 /lab_host="SOLR"
 /note="Vector: pBluescript SK(-); Site 1: EcoRI; Site 2:
 XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
 taken from tubers. The tubers were incubated at 26C in the
 dark for 2-3 weeks prior to sprouting. The eyes were
 frozen in liquid nitrogen immediately upon removal from
 tubers."

BASE COUNT 178 a 163 c 112 g 206 t
 ORIGIN
 Query Match 7.6%; Score 112.4; DB 155; Length 659;
 Best Local Similarity 53.0%; Pred. No. 2.3e-18;

Matches 287; Conservative 0; Mismatches 251; Indels 4; Gaps 2;
 Qy 27 CCGCGCCGGAGCTCGGTGGGAGCAATCGGTGGCGCTCACATTTCTTCGACATGAGCTGG 86
 Db 113 CTCCCGGAGGGGGCCAGAGCTTAACGCTCCCTCTTAACCTATTTTGTATCT-ATATGG 171
 Qy 87 CTGCAATTTTATCCCATGCTTCAGCTCTCTTCTAGCAATTCCTTTGTTCCAAAGCAACAT 146
 Db 172 TTCCGTTTGGGTATATGCTGGATTTTATCTACAAGCTACCAATTTTCCAAACCTCAT 231
 Qy 147 TTTTCAGAAATCGTTCCTCAAACTCAAACTCTCTCTAAAACCTCTCATACACTTC 206
 Db 232 TTGTTTCAAAACCATTTATCTCTTAAACATTTCACTCTCCCTCACTCTCAAAACACTAC 291
 Qy 207 TTCCCTCTCTCATGCAATTTTAAATCTACCTTCAATCCCGGAGAAATGCGGAGTTTCGG 266
 Db 292 ACACCTTTAGCTGGGCAACATGCTTGTCCACTAAATTCAGTGGTTATCTGAGTTGCGT 351
 Qy 267 TATCTATCCGGGAGCTCGGTTCTTTTACCATGCGAGAATCTAGCGAGCTTTCGATGAT 326
 Db 352 TATGTGACTGGAGATCTATATCTGTTACTTTTATTGAGACTGATATGAATTTCAATCAT 411
 Qy 327 CTGCTCGGAAATCGTCCGAAATCTCCGTTAGGCTCTACAACCTTTGTCCTTAATTTGCCG 386
 Db 412 CTCATTTGGTAACCATCTCTGNAATGCTAAGGATTTTATCCCTTCATTTCTCAATTTGGCA 471
 Qy 387 CCATTTGTGCAAGAAATCCGATAGAAACACTCTTCCAAAGTTTTCGCGTGCAGTGACTCTT 446
 Db 472 CAACCTAAGATGACACGGGGTCAAACTAGTCCCGGTCTTACCCATTTCAAGTGACACTT 531
 Qy 447 TTCCCGGAGCGAGCGCT--CGGTATTGGAATAGCAACGCATCACACCGTTAGCGAGGCC 503
 Db 532 TTTCGCAATCTTGGCATATCCCGTTGGTGTATTAGTAAACCATCATCTCGCTTGTGATGA 591
 Qy 504 CGCTCTTCTTCGCGTTTATACGGCTTGTGCTTCAATGAGCAAAACATTTGAAATGAA 563
 Db 592 AATACCATCGTGAATTTCAATAGAACATGGGTTTACTCAACAAATTCGCGGTGATGAA 651
 Qy 564 GA 565
 Db 652 CA 653

RESULT 11
 BE187619 788 bp mRNA EST 22-JUN-2000
 LOCUS EST336180 KV0 Medicago truncatula cDNA clone pkV0-1613, mRNA
 DEFINITION sequence.
 ACCESSION BE187619
 VERSION BE187619.1 GI:8666803
 KEYWORDS EST.
 SOURCE barrel medic.
 ORGANISM Medicago truncatula
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
 REFERENCE 1 (bases 1 to 788)
 VandenBosch K., Endre G., Hur J., Moore J., Beremand P., Ellis L.,
 Town C.D., Bowman C.L., Craven M.B., Hansen T.S., Holt I.E. and
 Fraser C.M.
 ESTs from uninoculated seedling roots of Medicago truncatula
 Unpublished (1999)
 CONTACT: VandenBosch K
 Department of Biology
 Texas A&M University
 College Station, TX 77843-3258, USA
 Tel: 409 845 7707
 Fax: 409 845 2891
 Email: kate@mail.bio.tamu.edu
 Texas A&M University name: T260795e
 TIGR sequence name: MTGAU50TK
 More information is available at:
 http://chrysis.tamu.edu/medicago

```
FEATURES
  source
    Seq primer: SKmod (CTA GAA CTA gta gAT CC) .
    Location/Qualifiers
      1..788
        /organism="Medicago truncatula"
        /cultivar="genotype A17"
        /db_xref="taxon:3880"
        /clone="pkV0-1613"
        /clone_lib="KV0"
        /tissue_type="Seedling roots"
        /dev_stage="Immediately prior to inoculation with
        Sinorhizobium meliloti (0 hour)"
        /lab_host="E.coli strain XL0LR"
        /note="Vector: pBluescript SK-; Site.1: EcoRI; Site.2:
        xhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
        was directionally ligated into the Unizap XR vector from
        Stratagene and packaged using Gigapack III Gold packaging
        extracts. Plasmids containing cDNA inserts were excised
        from the recombinant lambda-Zap phage using Ex-assist
        helper phage and propagated in XL0LR cells."
        237 a 188 c 100 g 262 t 1 others
BASE COUNT
ORIGIN
  374: Score 109.8; DB 164; Length 788;
  Best Local Similarity 51.18; Pred. No. 1.le-17;
  Matches 311; Conservative 0; Mismatches 292; Indels 6; Gaps 2;
QY 54 TCGGTGCGCTCACATCTCTCGCATGACGTGGTGCATTTTCATCCCATGCTTCAGCTC 113
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
5 TCCCTCCCCCAACATCTTTGACATTCCTTGGCTCTTTCTCTCCAGCAACCATCA 64
QY 114 CTCCTTCTAGCAATCCCTTGTTCGAAGCAACATTTTTCAGAAATCCATCGTTCCAAAATC 173
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
65 TTTTCTATGAATTTCTTCATCAATTTCCGATTTTCACACACCACCATTTGCCAAAAT 124
QY 174 AAACAATCTCTCTAAACATCTCATACACTCTCTCCCTCTCTCATGCAATTTAATCTAC 233
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
125 AAACAATCTCTCTCTCACACTTCAACATTAATCTCCCATTTTC---CGGTACATTCGTT 181
QY 234 CTTTCATCCCGGAGAAATGCGGAGTTTGGGTATCTATCTCGGGGACTCGGTTCTTTTC 293
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
182 CCTCGCTTGATCTAACCGAACCTCACTCGAATTTACTCTTAACAACCTCGTTTCATTC 241
QY 294 ACCATCGCAGAACTPAGCGAGACTTCGATGATCTGTCGGAATTCGCCAGAAATCTCCC 353
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
242 ACGGTGCGAGAGTCAATAGGAGACTTTGAACACTTGTCTCGGATTACTCAAGAGATGTT 301
QY 354 GTTAGGCTCTACAACCTTGTCCCTAAATTTGCGGCCCATTTGTCGAAGATTCGATAGAAA 413
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
302 AATGAATTCACCCCTTTGGTACCAAGTTGCAACAAA---TCTTTTCATTTTGAAGTAAAA 358
QY 414 CTCTCCAAAGTTTTCGCGGTGCAGGTGACTCTTTTCCAGCGCGGTCGGTATTGGA 473
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
359 GAGTTCCTTTGTTAGCTATTCAATATACATCATCTCCCAAAATATGCTTTTCCATTTGGC 418
QY 474 ATAGCAAGCATATCACACCGTTAGCGACGCCCGCTGTTCTCGGTTTATACGCGTTGG 533
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
419 CTAGCTTTTCATCATGTTGTTGCTGATGGAAGAACCTTCCACAAATTTTCATCAAAACATGG 478
QY 534 TCTTCATGAGCAACACATTTGAATGGAAGATGAAGATGAAGATTTAAATCTTTTGCCA 593
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
479 TCTTCATATTGTTCTCTCTCTACCTTTTGAAGACTCATCATCGCTTATAAAATCCCTTCCC 538
QY 594 GTTTTCGATAGATCGTCATAAAATATCCGACGAAATTTGACTCCATTTATTGGAGAAC 653
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
539 TTGATGATCGAAGTGTGATTATCGACACAATGACCTTCACGAAGTTTCTTGAAGAT 598
QY 654 GCGCTAAA 662
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
599 TGGAGAAA 607
RESULT 12
+ . . .
```

```
BE922746
LOCUS EST426515 574 bp mRNA EST 02-OCT-2000
DEFINITION EST426515 potato leaves and petioles Solanum tuberosum cDNA clone
          CSFB21L12 5' sequence, mRNA sequence.
ACCESSION BE922746
VERSION BE922746.1 GI:10448822
KEYWORDS EST.
SOURCE potato.
          Solanum tuberosum
          Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
          Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
          Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 574)
AUTHORS Van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
          C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
          Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
          Generation of ESTs from potato leaves and petioles
          Unpublished (2000)
          Contact: Cathy Ronning
          The Institute for Genomic Research
          For clone request: please contact Research Genetics, Libraries
          Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
  source
    1..574
      /organism="Solanum tuberosum"
      /cultivar="Kennebec"
      /db_xref="taxon:4113"
      /clone="CSFB21L12"
      /clone_lib="potato leaves and petioles"
      /tissue_type="leaflets and petioles"
      /dev_stage="8 weeks old plants"
      /lab_host="SOLR"
      /note="Vector: pBluescript SK(-); Site.1: EcoRI; Site.2:
      xhoI; Tissue was supplied by Dr. Fry (Cornell University).
      Leaflets and petioles were isolated from 8 week old
      greenhouse grown plants. The plants were watered and
      fertilized freely. The tissue was immediately frozen in
      liquid nitrogen."
      162 a 123 c 106 g 183 t
BASE COUNT
ORIGIN
  7.38; Score 108; DB 142; Length 574;
  Best Local Similarity 52.28; Pred. No. 3e-17;
  Matches 290; Conservative 0; Mismatches 260; Indels 6; Gaps 2;
QY 92 TTTTCATCCCATGCTTCAGCTCCTCTCTAGCAATTCCTTGTTCCAAGCAACATTTTTC 151
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
3 TTTTGGCTATATGCGTGGGATTTTATTCTACAAGCTACCAATTTCCAAACTCGATTCGT 62
QY 152 AGAATCCATCGTTCCAAAACATCAAAACATCTCTCTAAAACCTCTCATACACTCTTCCC 211
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
63 TCAACCATATTATCTCTACTCTTAACATTTCACTCTCCCTCACTCTCAACACATACACACC 122
QY 212 TCTCTCATGCAATTAATCTACCTTTCATCCCGGAGAAAATGCGGAGTTTCGGTATCT 271
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
123 CTTAGCTGCAACATGCTTCTGCCACTAAATTCGAGTGGTTATCTCGAGTTGCGTTATGT 182
QY 272 ATCCGGGAGCTCGGTTCTTTTCACCATCGAGATCTAGCAGACACTTCGATGATCTCGT 331
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
183 GACTGGAGATCTATATCTGTACTTTTATTGAGACTGATATGAATTTCAATCATCTCAT 242
QY 332 CGGAAATGCTCCAGATCTCCGTTAGGCTCTACAACCTTTTCCTTAATTTGCGGCCCAT 391
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
243 TGGTAACCATCTCGAAT---GCTAAGGATTTTATCCCAATTTCTCAATTGGCACACACC 299
QY 392 TGTCCGAAGAATCCGATAGAAAACCTTTCCAAAGTTTTCGCCGTGCGAGTACTTTTCCC 451
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
300 TAAGGATGACACGGGGTCAAACTAGTCCCGGTCTTAGCCATTCAGAGTGACACTTTTTC 359
QY 452 AGGCCGAGGCGTCGGTATTGGAATAGCAACCATCACACACCGTTAGCGACGCCCGCTGTT 511
DB 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11 11
360 GAATCTTTGGCATATCCGTTGGTTTTTAGTAACCATCATGTCGCTTGTGATGGAATACCAT 419
```

```

QY 512 TCTCGGCTTTATAACGGCTTGGTCTTCAATGAGCAACACACATTGAAATGAAGATGAAGA 571
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 420 CGTGAATTCATAAGAACATGGGTTTACTCAACAATTCGGCGGTGATGAACAGTGCCTT 479

QY 572 TGAAGAATTTAAATCTTTCCAGTTTTCGATAGATCCGTCATAAAATATATCCGACGAAATTT 631
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 480 AGAGAA---TGAGTTTCATTCCTATTTATGATAGGTCCGTGATAAAAGACCCTTATGAACA 536

QY 632 TGAATCCATTTATTGG 647
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 537 AGGACGATTATATGG 552

RESULT 13
LOCUS BF096979 482 bp mRNA EST 19-OCT-2000
DEFINITION EST400516 tomato nutrient deficient roots Lycopersicon esculentum
cDNA clone cLEW18020 5' sequence, mRNA sequence.
ACCESSION BF096979
VERSION BF096979.1 GI:10902689
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 482)
AUTHORS van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
TITLE Generation of ESTs from tomato nutrient-deficient roots
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU.
FEATURES
source
1..482
/organism="Lycopersicon esculentum"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone_lib="tomato nutrient deficient roots"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOLR"
/note="Vector: pBluescriptSMCUadap; Site_1: 5' EcoRI;
Site_2: 3' XhoI; Roots were harvested from plants grown
under the following deficiencies/stresses: 10 mM Al, Zn,
P, K. Fe.N. mRNA was isolated from individual treatments.
Proportional aliquots of mRNA of each treatment were mixed
and used for library construction."
BASE COUNT 113 a 119 c 93 g 157 t
ORIGIN

Query Match 7.2%; Score 106.6; DB 144; Length 482;
Best Local Similarity 51.6%; Pred. No. 6.7e-17;
Matches 244; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

QY 15 TGTAGATTTGGCGCGCGGACTCGGTGGCGGAGCAATCGGTGCGCTCACATCTTTC 74
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4 TGTCAAGTTGGCCACCTCCCGCGCGCAACGAGGTGATACCTCCTTACTATTATT 63

QY 75 GACATGACGTGGTGCATTTTCATCCCATGCTTCACGCTCCTCTTCTACGAATTCCTTGT 134
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 64 GACCATGTTTGGTTAGGTTTCGCCCGTATGAGCGCGGATATATTATTACAAGCTCTCCATT 123

```

```

QY 135 TCCAAGCAACATTTTTCAGAAATCCATCGTTTCAAAAACATCAAAACATCTCTCTAAAACT 194
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 124 TTCAAAACCGAATTCGTTTCAAAACATTAATCTCTCTCTTAAAAAATTCATCTCTCCCTCACT 183

QY 195 CTCATACACTTCTTCCCTCTCTCATGCAATTTAATCTACCCCTTCATCCCGGAGAGAAAATG 254
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 184 CTCAAACACATATACGCCCTTAGCCGGAAACGTTGCTGTGTGTCCTACCTAGATACAAACGGATAT 243

QY 255 CCGGAGTTTCGGTATCTATATCCGGGACTCGGTTTCTTCCACCATCCGAGAACTACGGAC 314
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 244 CCGTGAATTCAGTTATGTGACAGGAGATCTGTGCTGTGTACTTTTTCGAGACTGATATG 303

QY 315 GACTTCGATGATCTCGTCGTAATCTCCAGAAATCTCCCGTTAGGCTCTACAACTTTGTG 374
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 304 AATTTCAATTAATCTCATTTGTCACCATCCGCGTAAGGCTTAAGGATTTTATCACTTTGTT 363

QY 375 CCTAAATCCCGCCCATCTTCGGAAGAAATCCGATAGAAAACCTCTTCCAAAGTTTTCGCCGTG 434
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 364 CCTAAGTTAGGGGAACCTTAAGGATGCACCGGGGTCCAACCTAGCCCGCTCTTAGCCATT 423

QY 435 CAGGTGACTCTTTTCCAGGCGCGGTGCGGTATTTGGGAATAGCAACGCATCA 487
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 424 CAGGTGACACTTTTTCGAACTCTTGTGTATCCATTGGTTTCTACTAACCATCA 476

RESULT 14
LOCUS BE434257 591 bp mRNA EST 24-JUL-2000
DEFINITION EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG15H6, mRNA sequence.
ACCESSION BE434257
VERSION BE434257.1 GI:9432100
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 591)
AUTHORS Alcalá,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
source
1..591
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG15H6"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSMCUadap; Site_1: EcoRI;
Site_2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 159 a 129 c 121 g 182 t
ORIGIN

```

```
Query Match      7.2%; Score 106.6; DB 167; Length 591;
Best Local Similarity 51.6%; Pred. No. 7e-17;
Matches 269; Conservative 0; Mismatches 249; Indels 3; Gaps 1;

Qy 102 ATGCTTACGCTCTCTTCTAGCAATTCCTTGTTCACAGCAACATTTTTCAGATCCATC 161
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 7 ATGAGCGGATATTATTATTACAAGCTCTCCCATTTTCAACCCGATTTTGGTTCAAAACATT 66
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 162 GTTCAAAACCTCAACATCTCTCTTAACACTCTCATACACATCTTCCCTCTCTCATGC 221
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 67 ATTCCTCTCTTAANAATTCACCTCTCCCTCACTCTCAACACATATACGCCCTTAGCCGGA 126
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 222 AATTTAATCTACCCCTTCATCCCGGAGAAAATGCGGAGTTTCGGTATCTATCCGGGGAC 281
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 127 AAGCTTGCTTGTCACATAGATACAAACGGATATCTTGAGTTACGTTATGTACAGGAGAT 186
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 282 TCGGTTCTTTACCATCGCAGATCTAGCAGCACTTCGATGATCTCGTCGGAATCGT 341
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 187 TCTGTCTCTGTACTTTTTCGAGACTGATATGAATTTCAATTTATCTCATTTGGTGACCAT 246
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 342 CCAGAACTCCCGTTAGGCTCTACAACTTTGTCCTAAATTTGCCGCCCATTTGTGAGAA 401
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 247 CCGCTAGGCTAAGGATTTTATCACCCTTTGTCTTAAGTTAGGGGAACCTTAAGGATGCA 306
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 402 TCCGATAGAAAACCTTCCCAAGTTTTCGCCGTGCAAGGTGACTCTTTTCCCGAGCGGAGGC 461
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 307 CCGGGGTCCCACTAGCCGCTCTTACCCATTCAGGTGACACTTTTTCGGAATCTTGGT 366
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 462 GTCGGTATTGGAATAGCAACGATCACACCGTTAGCAGCGCCCGCTGCTTTCTCGCGTTT 521
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 367 GTATCCATTTGTTTTCACTAACCATCATCTTGTGTGATGAGCTACTATAGCAGGGTTC 426
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 522 ATAACGGCTTGGCTTCAATGAGCAACACATTTGAAATGAAGATGAAGATGAAGATTT 581
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 427 ATTAAGCGGTGGGCTCTACTCCACAAATTCGGTGGACATGAA---CAATTCATTTCGAAT 483
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 582 AAATCTTTGCCAGTTTCGATAGATCCCTCATAAATATCC 622
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 484 GAGCTAATTCATTTATGATAGTGGTCCGTAGTAAAGACCC 524
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 15
AW616206 487 bp mRNA EST 24-MAR-2000
LOCUS EST307245 L. hirsutum trichome, Cornell University Lycopersicon
DEFINITION hirsutum cDNA clone cLHTID15 5', mRNA sequence.
ACCESSION AW616206.1 GI:7322240
VERSION EST.
KEYWORDS
SOURCE Lycopersicon hirsutum.
ORGANISM Lycopersicon hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 487)
AUTHORS van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
,F., Hansen,T., Craven,M.B., Bowman,C.L., Konning,C.M., Nierman,W.,
Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
Generation of ESTs from wild tomato (Lycopersicon hirsutum)
trichomes
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
FEATURES
source
location/Qualifiers
1..487
```

```
/organism="Lycopersicon hirsutum"
/db_xref="taxon:62890"
/clone_lib="L. hirsutum trichome, Cornell University"
/dev_stage="mixed stages"
/notes="Leaves of various stages were shaken in liquid
nitrogen, shearing off trichomes. This procedure yielded a
mixture of cells which is highly enriched for trichome
likely with minor contaminations of other types of leaf
cells."
```

BASE COUNT 116 a 117 c 91 g 163 t
ORIGIN

Query Match 7.1%; Score 104.8; DB 118; Length 487;
Best Local Similarity 51.3%; Pred. No. 2e-16;
Matches 244; Conservative 0; Mismatches 232; Indels 0; Gaps 0;

Qy 26 GCCGCGCGGACTCGGTGGCGGAGCAATCGGTCCGCTCACATTCCTTCGACATGACGTG 85
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 GCCACCTCCCAAGCGCGCAACGAGGTGCTTCTTACTTATTTGACCATGTTTG 60
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 86 GCTGATTTTCATCCCATGCTTTCAGTCTCTTCTAGCAATTCCTTGTTCGAAGCAACA 145
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 GTTAGGGTTTCACCTTATTAGGCGGATATTATTCTACAAGCTCTCCATTTCCAAATCCGA 120
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 146 TTTTTCAGAAATCCATGCTTCCAAACACTCAACAACTCTCTCTTAAACTCTCATACACTT 205
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 TTTCTGTTCAAAACATTAATTCCTCTCTTAAATAATTCACCTTTCCCTCACTCTCAAAACATA 180
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 206 CTTCCTCTCTCATGCAATTTAAATCTACCCCTCATCCCGGAGAAAATGCCGAGTTTCG 265
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 TATGCCCTTAGCCGGAACGTTGCTTGTCCACTAGATACAAACGGATATCTGAGTTAGC 240
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 266 GTATCTATCCGGGAGCTCGGTTTCTTCCACCATCGCAATCTAGCGAGACTTTCGATGA 325
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 TTATGTGACAGGAGATCTGTGCTGTGTTAAATTTTCTGAGACTGATATGGATTTCAATTA 300
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 326 TCTGTCGGAATCGTCCAGAAATCTCCCGTTAGGCTCTAGCAACTTTGTCCTTAATGCC 385
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 TCTCATTGGCGACCATCCGCTAATGCTAAGGATTTTATCATTGTTCTCAAGTTAGG 360
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 386 GCCCATTTGTCGAAGAATCCGATAGAAAACCTCTTCCAACTTTTCCGCGTGAGGTCACCT 445
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 GGAACCTAAGGATGCACCGGGTCCCAATTAGCCCGCTCTTAGCCATTCAAGTGACACT 420
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Qy 446 TTTCCCGAGCGGAGCGTGGTATTGGAATAGCAACGCAATCACACCGGTAGCGAGC 501
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 TTTTCCGAATCTTGTGTATCCATTGGTTTTCACCTAACCATCATGTTGTGTGTATG 476
||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

Search completed: November 5, 2001, 18:00:56
Job time: 14896 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:50:26 ; Search time 8904.87 Seconds
(without alignments)
2619.394 Million cell updates/sec

Title: US-08-894-356C-5
Perfect score: 1508
Sequence: 1 TGAACATTCGACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba1.*
2: gb_ba2.*
3: gb_ba3.*
4: gb_in1.*
5: gb_in2.*
6: gb_in3.*
7: gb_om.*
8: gb_ov.*
9: gb_pat1.*
10: gb_pat2.*
11: gb_ph.*
12: gb_pl1.*
13: gb_pl2.*
14: gb_pl3.*
15: gb_pl4.*
16: em_ba1.*
17: em_ba2.*
18: em_fun.*
19: em_htgo_hum.*
20: em_htgo_inv.*
21: em_htgo_rod.*
22: em_htg_hum1.*
23: em_htg_hum2.*
24: em_htg_hum3.*
25: em_htg_hum4.*
26: em_htg_hum5.*
27: em_htg_hum6.*
28: em_htg_hum7.*
29: em_htg_hum8.*
30: em_htg_inv1.*
31: em_htg_inv2.*
32: em_htg_other.*
33: em_htg_rod.*
34: em_hum1.*
35: em_hum2.*
36: em_hum3.*
37: em_hum4.*
38: em_hum5.*
39: em_hum6.*
40: em_hum7.*
41: em_in.*
42: em_om.*
43: em_or.*

44: em_ov.*
45: em_pat.*
46: em_ph.*
47: em_pl.*
48: em_ro.*
49: em_sts.*
50: em_sy.*
51: em_un.*
52: em_vl.*
53: gb_sts1.*
54: gb_sts2.*
55: gb_sts3.*
56: gb_sy.*
57: gb_un.*
58: gb_vl1.*
59: gb_vl2.*
60: gb_htg1.*
61: gb_htg2.*
62: gb_htg3.*
63: gb_htg4.*
64: gb_htg5.*
65: gb_htg6.*
66: gb_htg7.*
67: gb_htg8.*
68: gb_htg9.*
69: gb_htg10.*
70: gb_htg11.*
71: gb_htg12.*
72: gb_htg13.*
73: gb_htg14.*
74: gb_htg15.*
75: gb_htg16.*
76: gb_htg17.*
77: gb_htg18.*
78: gb_htg19.*
79: gb_htg20.*
80: gb_htg21.*
81: gb_htg22.*
82: gb_htg23.*
83: gb_htg24.*
84: gb_htg25.*
85: gb_pr1.*
86: gb_pr2.*
87: gb_pr3.*
88: gb_pr4.*
89: gb_pr5.*
90: gb_pr6.*
91: gb_pr7.*
92: gb_pr8.*
93: gb_pr9.*
94: gb_r01.*
95: gb_r02.*
96: gb_in4.*
97: gb_pr10.*
98: em_ba3.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Length	DE ID	Description
1	1501.6	99.6	1508	10	E12757
2	202.4	13.4	1679	12	AB010708
3	202.4	13.4	1703	10	AB010708
4	189.2	12.5	1476	12	E12753
5	185.8	12.3	1479	10	AB029340
6	138.2	9.2	1622	10	E12756
7	138.2	9.2	1622	10	E12754
8	114.2	7.6	101176	12	AB026494
					AC002560
					Genomic s

```

9      88.2  5.8 119914 12 AC003027
10     86   5.7 1648   9 AX025514
11     72.6  4.8 66237 12 AB016892
12     60   4.0 1141 10 AX083744
13     55.8  3.7 805   53 CNS06V71
14     53.6  3.6 7218 10 I66494
15     53.2  3.5 321003 84 PFMAL4P3
16     52.4  3.5 2357 14 BNA9609
17     51.4  3.4 12229 4 AE001377
18     51.4  3.4 20941 67 AC021959
19     50.4  3.3 82360 13 AP000606
20     50.4  3.3 136047 74 AC069470
21     50   3.3 905   5 AF081800
22     49.8  3.3 85690 12 AB028618
23     48.6  3.2 136047 74 AC069470
24     48.6  3.2 3676 93 HSM801343
25     48.2  3.2 1612 94 AF056194
26     48.2  3.2 178273 60 AC005308
27     47.8  3.2 470 89 AK026556
28     47.8  3.2 2085 96 PFTHIORED
29     47.8  3.2 41035 13 AF287134
30     47.6  3.2 1954 89 AK000630
31     47.6  3.2 110000 84 PFMAL4P1_0
32     47.2  3.1 1442 13 AF329673
33     47.2  3.1 2000 6 DDCYTCVII
34     47.2  3.1 2982 6 DDISGSPA
35     47.2  3.1 13684 4 AE001403
36     47.2  3.1 310779 60 AC005140
37     47   3.1 395   5 AF151795
38     47   3.1 1881 5 AF008300
39     47   3.1 196556 73 AC068542
40     46.8  3.1 1124 53 CNS073BM
41     46.6  3.1 1124 53 CNS073BM
42     46.6  3.1 1534 10 I33391
43     46.6  3.1 114226 60 AC004710
44     46.6  3.1 196490 60 AC005507
45     46.2  3.1 1718 89 AK025339

```

ALIGNMENTS

```

RESULT 1
E12757
LOCUS      E12757      1508 bp      DNA
DEFINITION Senecio cruentus mRNA for acyltransferase,,partial cds.
ACCESSION E12757
VERSION   E12757.1 GI:3251589
KEYWORDS  JP 1997070290-A/5.
SOURCE    unidentified.
ORGANISM  unclassified.
REFERENCE 1 (bases 1 to 1508)
AUTHORS   Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
            Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE     GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL   Patent: JP 1997070290-A 5 18-MAR-1997;
            SUNTORY LTD
COMMENT   OS Senecio cruentus
            PN JP 1997070290-A/5
            PD 18-MAR-1997
            PF 30-JAN-1996 JP 1996046534
            PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
            ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
            MASAHIRO.
            PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC
            C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
            C12R1:865),
            PC (C12N9/10,C12R1:19);
            CC strandedness: Double;
            FH topology: Linear;
            Key Location/Qualifiers
            key

```

```

FT source 1..1508
FT /organism='Senecio cruentus'
FT /tissue_type='petal'
FT /clone='PCAI48'
FT CDS 1..1367
FT /product='acyltransferase'.
FEATURES
    source
    Location/Qualifiers
    1..1508
    /organism="unidentified"
    /db_xref="taxon:32644" 477 t
BASE COUNT 442 a 293 c 296 g 477 t
ORIGIN
Query Match 99.68; Score 1501.6; DB 10; Length 1508;
Best Local Similarity 99.78; Pred. No. 0;
Matches 1504; Conservative 0; Mismatches 4; Indels 0; Gaps 0:
QY 1 TGAACATTCGGAACATGCCGGAATATCGGCCCTCGGGCACCACATCGGCATCGCTCGT 60
DB 1 TGAACATTCGGAACATGCCGGAATATCGGCCCTCGGGCACCACATCGGCATCGCTCGT 60
QY 61 TATCTCTTACTTCTTCGACATTTACTTGGCTACTTCTCCCTCCGGTCCACCATCTTTCT 120
DB 61 TATCTCTTACTTCTTCGACATTTACTTGGCTACTTCTCCCTCCGGTCCACCATCTTTCT 120
QY 121 TCTATGACTTTCCACATTTCTAAATCCCATTTTCATGGACACTATTGTTCCCAAGGCTAAAC 180
DB 121 TCTATGACTTTCCACATTTCTAAATCCCATTTTCATGGACACTATTGTTCCCAAGGCTAAAC 180
QY 181 AATCTTTATCGGTGCACCTTTCAACATTTTTCGGCTTTGCTAGTAATTTGATTTATTTTC 240
DB 181 AATCTTTATCGGTGCACCTTTCAACATTTTTCGGCTTTGCTAGTAATTTGATTTATTTTC 240
QY 241 CTAACACTGATGGTTCGGGTTTTTAATAAAACCAGAAATAAAACACAGTTCGAAGGTGATT 300
DB 241 CTAACACTGATGGTTCGGGTTTTTAATAAAACCAGAAATAAAACACAGTTCGAAGGTGATT 300
QY 301 CTGTTGGTGTACTTTTCGAGAATGTTGCTGTAGACTTTTAATAATTTGACAGAAATCATC 360
DB 301 CTGTTGGTGTACTTTTCGAGAATGTTGCTGTAGACTTTTAATAATTTGACAGAAATCATC 360
QY 361 CTCGAAATGTGAAACATTTTATCCACTTTGACCTTTCATTTGGGAAATGCAATCAAAATAT 420
DB 361 CTCGAAATGTGAAACATTTTATCCACTTTGACCTTTCATTTGGGAAATGCAATCAAAATAT 420
QY 421 GTGATTCGGTCACGGTCCCACCTTTTTCACCTTCAAGTGACGCTTTTTCGGGGCTCGGTA 480
DB 421 GTGATTCGGTCACGGTCCCACCTTTTTCACCTTCAAGTGACGCTTTTTCGGGGCTCGGTA 480
QY 481 TATCACTAGGAATGACGAATCATATAGCCTTTGGTGACGCTTAGCACGCGGTTCAACTTTT 540
DB 481 TATCACTAGGAATGACGAATCATATAGCCTTTGGTGACGCTTAGCACGCGGTTCAACTTTT 540
QY 541 TGAAGGGTGGACTTCGATTTATCAATCTCGGTGTAGATCGGCTCTTTTAAACGAAGAT 600
DB 541 TGAAGGGTGGACTTCGATTTATCAATCTCGGTGTAGATCGGCTCTTTTAAACGAAGAT 600
QY 601 CTCACCGGTTTTGATAGATTGATTAAACATCCACATTTAGATGAAATAGTTGAGAC 660
DB 601 CTCACCGGTTTTGATAGATTGATTAAACATCCACATTTAGATGAAATAGTTGAGAC 660
QY 661 ATCAAGGCTCGAAAGTTTTTATAAACCCTTCGAGCCTTTGTTGGTCCCACTGATAAGTTC 720
DB 661 ATCAAGGCTCGAAAGTTTTTATAAACCCTTCGAGCCTTTGTTGGTCCCACTGATAAGTTC 720
QY 721 GGTCACCGTTTGTGTTGACCGCACTAATATCAATCTACTAAAGAAAGAGTCTTTAAACC 780
DB 721 GGTCACCGTTTGTGTTGACCGCACTAATATCAATCTACTAAAGAAAGAGTCTTTAAACC 780
QY 781 AAGTCCCAAACTTGGAGTACATGTCATCTTTTACGGTAACCTTTGTTGATATATGGAGTT 840
DB 781 AAGTCCCAAACTTGGAGTACATGTCATCTTTTACGGTAACCTTTGTTGATATATGGAGTT 840

```


QY 841 GCATAGCAATCACTCGTAAATAAGGAGAAAGGCGAAGACGAGTTAGAACACT 900
|||||
Db 841 GCATAGCAATCACTCGTAAATAAGGAGAAAGGCGAAGACGAGTTAGAACACT 900
QY 901 TCATATACCAATGATTCGATCGCTGCTGATGACCAATTCGCCAGCCTACTTTG 960
Db 901 TCATATACCAATGATTCGATCGCTGCTGATGACCAATTCGCCAGCCTACTTTG 960
QY 961 GTAACGTGGTGCACCATGTCGCCGACCTTAAATAATGCTTTTGACTACGGAATG 1020
|||||
Db 961 GTAACGTGGTGCACCATGTCGCCGACCTTAAATAATGCTTTTGACTACGGAATG 1020
QY 1021 GGTATGCACCTTGGTGTAAAGTAATTTGAGAGCTATATGCAAAATGATATATAAGG 1080
|||||
Db 1021 GGTATGCACCTTGGTGTAAAGTAATTTGAGAGCTATATGCAAAATGATATATAAGG 1080
QY 1081 ACGGAATCTTGAAGATGCGCGGAGATGCGATGCACTTTTCATGATCCCGGTAGGAAGA 1140
Db 1081 ACGGAATCTTGAAGATGCGCGGAGATGCGATGCACTTTTCATGATCCCGGTAGGAAGA 1140
QY 1141 TTGGTGTCTGCTACACCTAAGCTCACTTCTACGACCTTTCATGATCCCGGTAGGAAGA 1200
|||||
Db 1141 TTGGTGTCTGCTACACCTAAGCTCACTTCTACGACCTTTCATGATCCCGGTAGGAAGA 1200
QY 1201 GCATAAGTATGAGACTGTTTCAATAGACTAATACGTCGATTTCTATAAATGCAAGCA 1260
|||||
Db 1201 GCATAAGTATGAGACTGTTTCAATAGACTAATACGTCGATTTCTATAAATGCAAGCA 1260
QY 1261 AAACATCAGCACAAGATCTTGAATTTGATGAGTCTACCGAGTATGCAAAATGAGGCGT 1320
Db 1261 AAACATCAGCACAAGATCTTGAATTTGATGAGTCTACCGAGTATGCAAAATGAGGCGT 1320
QY 1321 TTTCTAGCATCTTGTATGAAGATTTAGAGATCAAGTTTCATGATGATCATCGTCCCT 1380
Db 1321 TTTCTAGCATCTTGTATGAAGATTTAGAGATCAAGTTTCATGATGATCATCGTCCCT 1380
QY 1381 TTTTGTGTCATCAAGTTTCTGCTGCTTTTATGAGTTCGACCTGTTCTATTCTTTAAGTA 1440
Db 1381 TTTTGTGTCATCAAGTTTCTGCTGCTTTTATGAGTTCGACCTGTTCTATTCTTTAAGTA 1440
QY 1441 TACCTTTCGACTATGTTTGAAGATGCAACGATATAAATGAAAAAAGAAAAA 1500
Db 1441 TACCTTTCGACTATGTTTGAAGATGCAACGATATAAATGAAAAAAGAAAAA 1500
QY 1501 AAAAAA 1508
Db 1501 AAAAAA 1508

RESULT 2
LOCUS AB010708
DEFINITION Gentiana triflora mRNA for Anthocyanin 5-aromatic acyltransferase,
complete cds.
ACCESSION AB010708
VERSION AB010708.1 GI:4185598
KEYWORDS Anthocyanin 5-aromatic acyltransferase.
SOURCE Gentiana triflora petal cDNA to mRNA, clone: pGAT4.
ORGANISM Gentiana triflora
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
1 (bases 1 to 1679)
REFERENCE
AUTHORS Fujiwara, H., Tanaka, Y. and Kusumi, T.
TITLE Direct Submission
JOURNAL Submitted (22-JAN-1998) to the DDBJ/EMBL/GenBank databases.
Hirotsuki Fujiwara, Suntory LTD., Inst. of Fundamental Research;
1-1-1, Wakayamadai, Yamamoto-cho, Osaka 618-0024, Japan
(E-mail: Hirotsuki.Fujiwara@suntory.co.jp, Tel: +81-75-962-8807,
Fax: +81-75-962-8262)
REFERENCE
AUTHORS Fujiwara, H., Tanaka, Y., Yonekura-Sakakibara, K.,

TITLE
JOURNAL
MEDLINE
FEATURES
source

Fukuchi-Mizutani, M., Nakao, M., Fukui, Y., Yamaguchi, M., Ashikari, T.
and Kusumi, T.
cDNA cloning, gene expression and subcellular localization of
anthocyanin 5-aromatic acyltransferase from Gentiana triflora
Plant J. 16 (4), 421-431 (1998)

Location/Qualifiers
1. .1679

/organism="Gentiana triflora"
/db_xref="taxon:55190"
/clone="pGAT4"
/tissue_type="petal"
6. .1415

CDS

/codon_start=1
/product="Anthocyanin 5-aromatic acyltransferase"
/protein_id="BA074428.1"
/db_xref="GI:4185598"
/translation="MEIQIMVKVLEKQVTPSPDTPVELSLPVTFFDPLHLNKM
SLLFYDFPPTDFLDVTPNLKASLSLTHYVPLSGNLLMLPIKSGMPKQFQYRDE
GDSITLIVAESQDFDLKGLHQLVDSNDLHGLFYVMPRTVMQYKVIPLVAVQTV
FNRGIAVALTAHHSIADAKSFVMPINAWYINKFGKADLLSANLLSFDRSIIKDL
YGLETFWNEQDVLMEFMRSGSKPPRENKRYATYVLSLAELQKLNKLVNLRGSEPT
IRVITFTTTCGTWTCMKSDVVSEESNDNELEYFSETADCGLTGTPCPNPF
GNCLASCAVAKATHEKLVGDLVAVAAIGEAIEKRLHNEKGVLDADKTWSESNGIP
SKRFLGTCSPKFDSDYGVDFGKPAKFDITSDVYAEIYVQSRDFEKGVEIGVSLP
KIHMDAPAKIFEFGFCSLS"
1679

polya_site

/note="24 a nucleotides"

BASE COUNT 488 a 350 c 359 g 482 t

ORIGIN

Query Match 13.4%; Score 202.4; DB 12; Length 1679;
Best Local Similarity 52.0%; Pred. No. 1.5e-34;
Matches 720; Conservative 0; Mismatches 616; Indels 48; Gaps 10;

QY 1 TGAACATCTCTCAACATGCGGCAATATCGGCCCTCGGGCACCACCA---TCGGCCATCGCT 57
|||||
Db 25 TGAAGTCTCTGAAAATGCGCAAGTTACACCACCATCTGCACACACAGATGTCGAGTTAT 84
QY 58 CGTTATCTCTTACTTCTTGACATTTGCGTACTTTCCTCGGTCGCGGTCACCATCTTTT 117
|||||
Db 85 CGTACCGGTAAACATCTTTCGATATCCCTGTTGCACTTGAATAAGATGCACTCCCTTC 144
QY 118 TCTTCTATCATCTTCCACATTTCTAATCCATTTCTGACGACACTATTGTCGCCAGCTAA 177
|||||
Db 145 TGTTTTACGACTTTTCGTTACCCCAAGAACACATTTCTTTGGACACTGTTATCCCTTAATCTTA 204
QY 178 AACAACTTTTATCGGTCACTCTTCAACATTTTTCCTCGGTTTGTAGTAATTTGATTGTAT 237
|||||
Db 205 AGGCCTCTTGTCTCTCACTCTAAACACACTAGCTTCCGCTTAGCGGAATTTGTTGATGC 264
QY 238 TTCCCTAACACTGATGGTTTCGGGTTTAAATAAAACACAGAAATAAACACGTTGAAGGTG 297
|||||
Db 265 CGATCAA-----ATCGGCGAAATGCCAAGTTTTCAGTACTCCCGTGATGAGGGCG 315
QY 298 ATCTGTTGTTGTTACTTTTTCGAGAATGTTGCTTCACTTTAATAATTTTCACAGGAATC 357
|||||
Db 316 ACTCGATAACTTTTGATGCTGCGGAGTCTGACCAAGGATTTTGACTACCTTAAAGGATC 375
QY 358 ATCCCTCGAAATGTGAAAACTTTTATCCACTTGTACTCTTTCATTTGGGAAATGCAATCAAT 417
|||||
Db 376 AACTGTTAGATTTCCATGATTTGCATGGCTTTTATGTTATGTCACCGGTTATTAAGGA 435
QY 418 ----TATGTTGATGCGTCACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTTTCGGGGCT 474
|||||
Db 436 CCATGCAAGACATATAAGTGATCCCGCTGTCAGCCGTCGAAGTAACCGTTTTTTCCTAACCC 495
QY 475 CGGGTATATCACTAGTAAGTACCAATATCATCATAGCCCTTGTGACGCTACGACCGGGTTCA 534
|||||
Db 496 GTGGCATGCCCGTGGCTCTGACGGGCACATCATCAATTCAGATGCTAAAGGTTTGTAA 555
QY 535 ACTTTTGAAGGGTGGACTTCGATTATTCAATCTGTTGATAGATCGTCTTTTTTAAACGA 594

Db	556	TGTTCAATGCTTGGGCTATATTAAACAATTTGGGAAGACCGGACTGTGTGTCG	615
Qy	595	AAGGATCTCCACCGGTTTGTATAGATTGATTAAACA-----TCCACATTTTAGATGA	646
Db	616	CGAATCTTCTCCATCTTTCGATAGATCGATAATCAAGAATCTGTATGGCCTAGAGGAA	675
Qy	647	AAATAAGTTGAGACATACAAG----GCTCGAAAGTTTATTAACCTTCGAGCCCTGTTG	702
Db	676	CATTTTGGAAACGAATGCAAGATGTTCTTGAATGTTCTCTAGATTTTGGAAGCAACCC	735
Qy	703	GTCCACTGATAAGTTCGGTCAAGCTTTGCTTGACCGCAACTAAATCAATCTACTATA	762
Db	736	CTCGATTCAACAGGTACGAGCTACATATGCTCTCCCTTGTCTGAATCCAGAAGCTAA	795
Qy	763	AGAAAAGGTCCTTAA----CCCAAGTGCCAAAACCTTGGAGTACATGT--CATCTTTTACGG	816
Db	796	AGACAAAGTACTGAATCTCAGAGGATCCGAACCCACAAATACGTGTAAACGAGTTCACA	855
Qy	817	TAAGTTGTGGTTATATATGAGAGTTGATAGCGAAATCA-----CTCGTAAAAATAG	867
Db	856	TGACGTGTGATACGTATGACATGATGCGTCTCAAAATCAAAAGATGACGCTGTATCAGAGG	915
Qy	868	GAGNAAGAAAGGCGAAGCAGGATTAGACAGTTTACAATCACCATGATGATGTCGATCTC	927
Db	916	AATCATCGAACACGAGAAATGAGCTCGAGTACTTCAGTTTTCACAGCGGATTCGCGAGGAC	975
Qy	928	GTCTTGATCCACCAATTCCCACAGGCTACTTTGGTAACTGGTGGCACCATGTCTCCCGA	987
Db	976	TTCTGACGCCCGTCGCCCTAAGTCTTGGCACTGTCTTGGCGTCATGCGCTTGCAG	1035
Qy	988	CCTTAAAAATGCTGTTTGGACTACGGAATATGGTATGCACCTTGGTCTAAAGTAATG	1047
Db	1036	AAGCAACACATAAAGAGTTAGTTGGGATAAAGGCTTCTTTGTTGCAGTTGCAGCTATTG	1095
Qy	1048	GAGAGCTATATGCAAAATGATATATAAGGACGGAACTCTTGAAGATGCCCCGAGAT	1107
Db	1096	GAGAAGCCATTGAAAAGAGGTTGCAACACGAAAAAGGCGTTCTTTCGAGATGCAAAAACTT	1155
Qy	1108	GG---CATGACACCTTCATGATCCCGGCTAGGAAGATT---GGTGTGCTGGTACACCTA	1161
Db	1156	GTTTATCGGAATCTATGGAATCCCTTCAAAAGATTTCTCGGATTTACCGGATCGCCTA	1215
Qy	1162	AGCTCAACTTGTACGACTTTGATTTGGGTGGGGAAGCGCATAAAGTATGAGACTGTTT	1221
Db	1216	AGTTCGATTCGTATGGTGTAGATTTTGGATGGGAAAGCCTGCAAAATTTTGACATTACCT	1275
Qy	1222	CANTAGCTATATACGTCGATTTCTATTAATGCAAGCAAAACATCAGCACAAAGCTTG	1281
Db	1276	CTGTTGATTATGCAGAATTGATTATGTGATTTCAGTCCAGGAGTTTGGAAAAAGGTGG	1335
Qy	1282	AAATTGGATTGAGTCTACCGAGTATGCAAAATGGAGCGCTTTCTTAGCATCTTTGATCAAG	1341
Db	1336	AGATTGGAGTATCATTTGCCCTAAAGTTTCATATGGATGGATTTGCAAAAATCTTTGAAGAAG	1395
Qy	1342	GATT 1345	
Db	1396	GCTT 1399	
RESULT	3		
LOCUS	E12753	1703 bp	DNA
DEFINITION	Gentianatriflora mRNA	acyltransferase, complete cds.	24-JUN-1998
ACCESSION	E12753		
VERSION	E12753.1	GI:3251585	
KEYWORDS	JP 1997070290-A/1.		
SOURCE	unidentified.		
ORGANISM	unclassified.		
REFERENCE	1 (bases 1 to 1703)		
AUTHORS	Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Ybnekura, K., Mizutani, M. and Kusumi, T.		

[illegible]

Db 496 GTGGCATAGCCGCTGCTGACGGCACATCATTTCAATTCAGATGCTAAAGTTTTGTAA 555
QY 535 ACATTTTGAAGGGTGCAGCTTCGATTATTCATCTGCTGTAGATCGTCTTTTAAACGA 594
Db 556 TGTTCATCAATGCTGGGCTATATTACAAATTTGGGAAGACGGGACTGTTGTCGG 615
QY 595 AAGGATCTCCACCGGTTTGTAGATGATTAAACA-----TCCACATTTATGATGA 646
Db 616 CGAATCTTCTCCATCTTTCGATAGATCGATATCAAGATCTGTATGCCGTAGAGGAA 675
QY 647 AATAAGTTGAGACATACAAG-----CCTCGAAGTTTATTAACCTTCGAGCCTTGTG 702
Db 676 CATTTTGGAAACGAATGCAAGTGTCTTGAATGTTCTCTAGATTGGAAGCAAAACCC 735
QY 703 GTCCCACTGATAAAGTTCGGTCAACCTTTGTGTGACCGCACTATATCAATCTACTAA 762
Db 736 CTGATTCACAAGGATGACGCTACATATGTCTCTCCCTTGTCTGAAATCCAGAGCTAA 795
QY 763 AGAAAAGGCTTTAA-----CCCAAGTGCACAACTTGGAGTACATGT--CATCTTTTACGG 816
Db 796 AGAACAAGTACTGAATCTCAGAGGATCCGACCGACATACGTGTAAACGACGTTCA 855
QY 817 TAACCTTGTGTTATATATGAGATGTCGATACGGAATCA-----CTCGTAAATAAG 867
Db 856 TCACGCTGTGGATACGTATGGACATGCTGCTCAATCAAAAGATGACGTGATCAGAG 915
QY 868 GAGAAAAGGCGGAGGAGTGTAGACAGTTCATATCAATCACCATTGATGTCGATCTC 927
Db 916 ATATCTGACGACGAGAAATGAGCTCGAGTACTTCAAGTTTACAGCGGATTCGCGAGG 975
QY 928 GTCTTGATCCACCAATTTCCACAGCCTACTTTGGTAAGTGTGTGACCATGTCGCCGA 987
Db 976 TTCTGACGCCCCGCTCGCCCTAACTACTTTGGCAACTGTCTTGCATGCGTTGCAA 1035
QY 988 CCTTAAAAATGCTGTTTGTGATACGGAATGCGGTATGCACTTGTGCTAAAGTATTG 1047
Db 1036 AAGCAACACATAAAGAGTTAGTTGGGGATAAAGGGCTTCTTGTGCGATTTGCACT 1095
QY 1048 GAGAGCTATATGCAAAATGATATATAAAGGACGGAATCTTTGAAAGATGCCCGGAG 1107
Db 1096 GAGAACCCATTTGAAAGAGTTGCAACAGCAAAAGCGCTTCTGAGATGCAAAACTT 1155
QY 1108 GG---CATGAACCTTTTCATGATCCCGCTAGGAAGATT---GGTGTGCTGTACACCTA 1161
Db 1156 GGTATCGGAATCTAATGGAATCCCTTCAAAAGATTTCCTCGGATACCGGATCGCTA 1215
QY 1162 AGCTCAACTGTGACGACTTTGATTTGGGTGGGGAAGCGCATAAAGTATGAGACTGTT 1221
Db 1216 AGTTCGATTGCTATGCTGATGATTTTGGATGGGGAAGCGCTGCAAAATTTGACATT 1275
QY 1222 CAATGACTATATACGTCGATTTCTATAAATGCAAGCAAAACATCAGCACAGATCTTG 1281
Db 1276 CTGTTGATTATGAGAAATGATTATATGATTCAGTCCAGGGATTTGAAAGAGGTGG 1335
QY 1282 AATTTGATTGAGTCTACCGAGTATGCAATGAGGCGTTTCTAGCATCTTTTGTATGAAG 1341
Db 1336 AGATTGGAGTATCATTTGCCTAAGATTCATATGATGATGCAATTTGCAAAATCTTT 1395
QY 1342 GATT 1345
Db 1396 GCTT 1399

RESULT 4

AB029340
LOCUS AB029340 1476 bp mRNA PLN 04-APR-2000
DEFINITION Perilla frutescens mRNA for anthocyanin acyltransferase, partial cds.
ACCESSION AB029340
VERSION AB029340.1 GI:7415645
KEYWORDS anthocyanin acyltransferase.
SOURCE Perilla frutescens leaf cDNA to mRNA.
ORGANISM Perilla frutescens

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

JOURNAL

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
1 (sites)
Sakakibara, K.Y., Panaka, Y., Mizutani, M.F., Fujiwara, H., Fukui, Y., Ashikari, T., Yamaguchi, M. and Kusumi, T.
Molecular and biochemical characterization of a novel hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6-unpublished (1999)
2 (bases 1 to 1476)
Sakakibara, K.Y. and Tanaka, Y.
Direct Submission
Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Keiko Y Sakakibara, Sunito Research Center, Fundamental Research, Plant Biotech: Shimamoto-cho, Wakayamadai, 1-1-1, Mishima-gun, Osaka 618-8503, Japan (E-mail: Keiko_Sakakibara@suntory.co.jp, Tel:+81-75-962-8807, Fax:+81-75-962-8262)
Location/Qualifiers
1. .1476
/organism="Perilla frutescens"
/db_xref="taxon:48386"
/tissue_type="leaf"
<1. .1343
/codon_start=3
/product="anthocyanin acyltransferase"
/protein_id="BA093475.1"
/db_xref="GI:7415646"
/translation="VETCRVGPDPDSVAEQSVPLTFEDMTLWHPMLQLLFYEPFC SKDHFSEIVPKLQSLKTLIHFFPLSNLIYSPSPKPEFYLFPSGVSFTTAIES SDOFDLVGNRPESVRLYFVPLPIVEEDSRKLFQVFAVOTLFGDGVGIGIAT HHTVSDAPSEFLAITAMSSMSKHIEDDEEFKSLPFDVSRVIRYPTKFDISYWRNA LKFLQSRPSLPTDRITFTVFTQSKIKKLGWITQSRVPSLVHLSFVAIAAYMMAG ITKSTADEDONEDAFPLIPVDLRPLDPVPENYFGNCLVALPRMRRLVGEKG VFLEAAETAEIKKRINDKRILETVEKSPERKALOKSYFVAGSSKLDLXGADFGW GKAKGQELISDGEYAMTLCARDFEGGLEVCLSLPKDKMDAFRAFYSLGING"

BASE COUNT
ORIGIN

Query Match 12.5%; Score 189.2; DB 12; Length 1476;
Best Local Similarity 48.5%; Pred. No. 1.2e-31;
Matches 725; Conservative 0; Mismatches 33; Indels 33; Gaps 6;

QY 5 CATTCGCAACATCGCCGCAATATCGGCCCTCGGGCACCATCGGCATCGCTCGTTATC 64

Db 2 CGTGATCGAAACGTGTAGAGTTGGCGCGCGCGGACTCGGTGGCGGAGCAATCGGTGCC 61

QY 65 TCTTACTTCTTCGACATTAATTTGGGTACTTTCCTCCGCTCCACCATCTTTTCTTCTA 124

Db 62 GCTCACATCTTCGACATGACGTGGGTGCAATTTTCATCCCATGCTTCAGCTCTCTTCTA 121

QY 125 TGACTTTCACATTTCAATTCCTATTTTCATGGACATATTTGCCAGGCTAAACAATC 184

Db 122 CGAATTCCTTGTTCGAAGCAACATTTTTCAGAATCCATCTGTTCCAAACATCAACAATC 181

QY 185 TTTATCGGTCACTCTCAACATTTTTCCTGTTTGGTGTAGTAATTTGATTTCTTCTAA 244

Db 182, TCTCTCTAAACATCTCATACCTTCTCCCTCTCTCATGCAATTTAATCTACCTTCATC 241

QY 245 CACTGATGTTTCGGGTTTTTAATAAAAAACAGAAATAAAACACGTTTGAAGGTGATCTGT 304

Db 242 CCGCG-----AGAAATGCGGAGTTTCGGTATCTATCCGGGACTCGGT 286

QY 305 TGTGGTTACTTTTGCAGATGTTGTCTTGACTTTTAATAATTTGACAGGAATCATCTCG 364

Db 287 TTCTTTTCCACATCGCAAGATTCAGCGACGACTTCGATGATCTCTCGGAAATCTCCAGA 346

QY 365 AAAATGTGAAACATTTTATCCACTTGTACCTTCATTTGGGAAATGCAATCAAAATATGTGA 424

Db 347 ATCTCCGTTAGGCTCTACAACTTTTGTCCCTAAATTCGCCGCCCATTTGTCGAAGAAATCCGA 406

QY 425 TTGCGTCACGCTCCCACTTTTTCACCTTCAAGTGACGTTTTTTCGGGCTCGGATATC 484


```
Db 347 ATCTCCGTTAGGCTCTACAACATTTGTCCCTAAATTTGCCGCCCATTTGTCGAAGAATCCGA 406
QY 425 TTGGGTACCGTCCCACTTTTTCACCTTCAAGTGACGTTTTTTCGGGGCTCGGTATATC 484
Db 407 TAGAAAACCTCTTCCAAGTTTTTCCGCTGCGAGTGACTCTTTTCCAGCGCGAGGCTCGG 466
QY 485 ACTAGGAATGACCAATCATAGCTTGTGTAGCTAGCACCGGTTCACATTTTTTGAA 544
Db 467 TATTGAATAGCAACGATCACACCGTTAGCGAGCGCCGCTGTTCTCGCTTTATAAC 526
QY 545 AGGGTGACTTCGATTAATCAATCTCGGTGTAGATCGGTCTTTTTTAACGAAGG---ATC 601
Db 527 GGCTTGGTCTTCAATGAGCAACACATTCGAAATGAAGATGAAGATCAAGAATTTAAATC 586
QY 602 TCCACCGGTTTTTGATAGAT---TGATTACATCCACATTTTGTAGATCAAAATTAAGTTGAG 658
Db 587 TTTGCCAGTTTTCGATAGATCCGTCATAAATATCCAGCAAAATTTGACTCTCACTTTATTG 646
QY 659 ACATACAAGGCTCGAAAGTTTTTATAAACCTTCGAGCGCTTTGTGGTCCCACTGATAAAGT 718
Db 647 GAGAACGCGCTAAATTTCTTTTGCATCTCGTCATCTCCCTCATACCGAGCGACCGCAT 706
QY 719 TCGGTCAACGTTTGTGTGACCGAATATATCAATCTACTAAAGAAAAGGTCTTTAAC 778
Db 707 TCGAACCCAGCTTGTGTTTCAACCAATCCAAATTAAGAAATTCGAAGGTTTGGATTGAGTC 766
QY 779 CCAAGTCCCAAACTTGAGTACATCTCTTTTACGGTAACCTTGTGGTTATATATGGAG 838
Db 767 CAGATTTCGAAGTTTGTAGTCCATCTCTCTTTTGTAGCGATTTGACGCTTATATGTGGC 826
QY 839 TTGCATAGCGAAATCACTCTGTAATAATAGGAGAAAGGCGGAGAGAGGTTAGAAC 898
Db 827 TGCATTAACGAATCATTC-----ACAGCAGATGAGACCAAGACACGAGGATGCATT 880
QY 899 GTTCATAATCACCATTGATGTGATCTCGTCTGTGATCCACCAATTTCCACAGCCTACTT 958
Db 881 TTTCTGATTTCGGTTCGATCTAAGGCCAGATTAGATCCGCCGCTTCTCGAAATTAATT 940
QY 959 TGCTAACTGTGGTGCCACCATGTCTCCGACCTTAAAAAATGTCGTTTGTACTACGGAAA 1018
Db 941 CGGGAACGTGCTTATCGTACGCGCTGCGGAGAAATGCGCGCGGAGAGCTGTGTGGAGAGAA 1000
QY 1019 TGGGTATGCACTTGGTGTAAAGTAAATTTGGAGAGTCTATATGCAAAATGATATATAATA 1078
Db 1001 AGGGTGTCTTCTGCGACTCAGCTAATCGCGCGGAGATAAAGAAAGGATCAACGACAA 1060
QY 1079 GGAGGGAATCTTGAAGATGCCGGGAGATGGCATGA---ACCTTTCAATGATCCCGGCTAG 1135
Db 1061 GAGAATATTAGAAACGGTGGAGAAATGGTCCCGGAGATTCGTAAGCGTTGCGAGAAATC 1120
QY 1136 GAAGATTGGTGTGCTGTACACTTAAGCTCAACTGTACGACTTTTGATTTTGGGTGGG 1195
Db 1121 ATATTTTTCGGTGGCAGGATCGAGCAAGCTAGATCTTTACGGTGCAGATTTTGGATGGG 1180
QY 1196 GAAGCGCATAAAGTATGAGACTGTTTCAATAGA---CTAATAACGTCGATTTCTATATAA 1252
Db 1181 GAAGCGGAGAAACGAATAATTTGCGATTGATGGGAGAAATATCAATGACRCTTTG 1240
QY 1253 TCGAAGCAAAACATCAGCAAGATCTTGAAATTTGGAATTTGAGTCTACCGAGTATGCAAT 1312
Db 1241 TAAAGCCAGGGAATTTCAAGGAGGATTTGGAGGTTTGCTTCTTGTGCTTAAGGACAAAT 1300
QY 1313 GGAGGCGTTTTCTAGCATCTTTGATGAAGGATTTAGAGAGTCAA 1355
Db 1301 GGATGCTTTTGTGCTTATTTTTCACCTGGGAATTAATGGTTAA 1343
```

RESULT 6

```
E12754 LOCUS E12754 1622 bp DNA PAT 24-JUN-1998
DEFINITION Gentianatriflora mRNA for acyltransferase, complete cds.
ACCESSION E12754
VERSION E12754.1 GI:3251586
```

```
KEYWORDS JP 1997070290-A/2.
SOURCE unidentified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1622)
AUTHORS Ashikari,T., Tanaka,Y., Fujiwara,H., Nakao,M., Fukui,Y.,
Yonekura,K., Mizutani,M. and Kusumi,T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 2 18-MAR-1997;
SUNTORY LTD
COMMENT OS Gentianatriflora
PN JP 1997070290-A/2
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO,
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,(C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH Key Location/Qualifiers
FT source 1..1622
FT /organism='Gentiana triflora' FT
FT /variety='japenica'
FT /tissue_type='petal'
FT /clone='pGAT106'
FT CDS 35..1474
FT /product='acyltransferase'.
FEATURES
Source 1..1622
/db_xref='taxon:32644'
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN
```

```
Query Match 9.2%; Score 138.2; DB 10; Length 1622;
Best Local Similarity 48.6%; Pred. No. 2e-20;
Matches 676; Conservative 0; Mismatches 643; Indels 72; Gaps 8;
```

```
QY 1 TGAACATCTCGAACATGCGCGAATATCGCCCCCTCGGCACCATCGGCATCGCTCGT 60
Db 57 TCAAAGTCTTTGAGAAATGCGGTGCGCCACACCGGACGCGTCCGCGAGTTTACAG 116
QY 61 TATCTCTTACTTTCTCGACATTAATCTGGCTACTCTTCCCTCGGGTCCACCATCTTTTCT 120
Db 117 TCCCACTGTCTTTTCGACATCGGATGGTGTGATCTCTGATCGAGAACACCATCTGCATT 176
QY 121 TCTATGACTTTCCACATCTCTAAATCCCATTTTCATGGACACTATTTGTTCCAGGCTAAAC 180
Db 177 TCTACAGATTCGCGCATCTCTTGTCCCAACTCTAAATTTATCATTTTATCCCA---TTAAAT 233
QY 181 AATCTTTATCGGTCACTCTTAAACATTTTTCGGTTTCTGCTAGTAATTTGATCTGATTTC 240
Db 234 CGTCCCTCTTCCCTTGTCTCAAAACACTTCTCTCCGTAGCGGGAATTTGATTTGGCCGG 293
QY 241 CTAACACTGATGGTTCGGGTTTTTAATAAAAAACAGAAATAAAAAACAGTTGGAAGTGATT 300
Db 294 TAGATTCCTCG-----ATAGATGCGGAGTTGCGTTACAGAAAGGGGACT 341
QY 301 CTGTTGTGTTACTTTTTCAGAAATGTTGCTTGTGACTTTAATAATTTGACAGAAATCATC 360
Db 342 CCCTTTCTTTAACAATTCGAGATCGAGCATGGATTTGATTTGATTTATCTCGCCGAGATCATC 401
QY 361 CTGAAATGTGAAACTTTTATCCACTTGTACCTTTCATTTGGGAAATGCAATCAATAT 420
Db 402 AGAGGGATCTTTAATAATTCACGATTTGATTCGCGAGCTCCCAAGACCGATTTGTAACCT 461
QY 421 GTGATTGCGTCACGGTCCCACTTTTTCCTACTTCAAGTGACGTTTTTTCGGGGTTCGGGTA 480
```

Db 462 CGGCGAGAGTATTACCACTTTTGGTTTACAGGTGACGGTGTCTCCACACCGGTA 521
Qy 481 TATCACTAGGAATACGAATCATATAGCTTGGTGGAGCTAGCACGGGTTCACACTTTT 540
Db 522 TATGATTTGGAGCGCAATCTTCATCAAGTCTCTGGTGATGCCAGTCTCTTTCGCAATTTA 581
Qy 541 TGAAGGGTGGACTTCGATATTTCAATCTGGTGTAGAT-----CGGT 582
Db 582 ATAAATTTATGGTCTTGGTTGACAAATCCAATGGAGATTTCATTAAGTTCTCTCCACTTT 641
Qy 593 CTTTTTTAACGAAAGATCTCCACCGGTTTTTGATAGATTCATTAACATCCACATTTAG 642
Db 642 CTTCTCTACCTATGACGACAGATCTGGTGGCAAGATCCATTTTCATATCGTCAAAAA 701
Qy 643 ATGAAATAAGTTGAGACATACAAGGCTCGAAAGTTTTTATAAACCTTCGAGCGTTGTG 702
Db 702 TCTACAATGAAGAANAATGCTCAATCTCAGGGCACACCTACTGTTCTAAATCCAGCAA 761
Qy 703 CTCCTACTGATAAAGTTCGGTCAAGCTTTGTGTTGACCCGAACCTAATCAATCTACTTAA 762
Db 762 TTTCTAAAGATGAAGTTCGAGCCACTTCATCTACACCTTATGATATCATGAAGCTCA 821
Qy 763 AGAAAAAGGCTTAACCAAGTGCCAAACTT-----GGAGTACATGTCAT 807
Db 822 AGAAATTCATTTTCGTCAAAAAATCGCAACTTAACCGGTAGTAGTAATTAATCTGTCAA 881
Qy 808 CTTTTACGGTAACTTGTGTTATATATGAGTTCATAGAGTTCATAGCGAATCACTCGTAAATAAG 867
Db 882 CTTTCAGGATGACATCTGCATGATCTGGAGATGCTTCTGCAAAATCTTTCGAAATCATTAGACACCGTCG 941
Qy 868 GAGAAAGAAAGGGCGAAGACG-----AGTTAGAACATGTTTCATATCACTACCATGATTTGTC 921
Db 942 TAAGAGAGAAGGTGGAAGAGATAAACATGACGACCAACTTATGCTTTTCATCACTGCC 1001
Qy 922 GATCTCGTTCGATCCCAATTCACAGCGCTACTTTGGTAACTGTGGTGACCATGTG 981
Db 1002 GACAACGTTTTGCTCCGCGATACCTCAAAATTAATTTTGGAAATGTCATAGTGCCTTGA 1061
Qy 982 TCCCGACCTTAAAAAATGCTGTTTGTACTACGGAATAATGGGTATGCTACCTTGTGCTAAAG 1041
Db 1062 TGGTGGGATCGACTCATGAGCAACTTGTAGGAATGAAGGGTTGCGGTAGCTGCAACCG 1121
Qy 1042 TAATTTGAGAGTCTATATGCAAAATGATATATAAAGGACGGAATCTTTGAAAGATGCG 1101
Db 1122 CCATCGAGATGCTATCCATAAGAGGTTTACATGACTACGAAGGAATCTCGAGAGGAGATT 1181
Qy 1102 CGAGATGCGATGAACCTTTCATGATCCCG-----CTAGGAAGATTGGTGTGCTG 1152
Db 1182 GGATATCGGCGCGCGATCAACATCTGCGGCAACCAAGGTCGACGCTCATTTATGTGCTGTG 1241
Qy 1153 GTACACCTTAAGCTCAACTTGTACGACTTT-----GATTTTGGTGGGGAAGCGCATAA 1206
Db 1242 GATCGGACACGCAATGTGATGATTTTGGTGAGATTTTGGTGGGGAAGCTTGAAA 1301
Qy 1207 AGTATGAGACTGTTTCAATAGA---CTATAATAGCTGATTTCTATATAATCAAGCAAAA 1263
Db 1302 AGCATGAATCTGTTTCAACTAATCTTCGCGCAACACATAATTTTGTCTCTCGGTCAGAA 1361
Qy 1264 CATAGCACAGATCTTGAATTTGGATTGAGTCTACCGAGTATGCAATGGAGGCGTTTT 1323
Db 1362 GATTTAAAGGAGCACTTGAGCTTGCAATTTCTTTGCTTAAAGATAGGATGAGCAGCTTTG 1421
Qy 1324 CTAGCATCTTT 1334
Db 1422 CCACCATTTTT 1432

RESULT 7
AB026494 1622 bp mRNA PLN 04-APR-2000
LOCUS AB026494 1622 bp mRNA for acyltransferase homolog, complete cds.
DEFINITION Gentiana triflora GAT106 mRNA

AB026494.1 GI:7415596
acyltransferase homolog.
Gentiana triflora cDNA to mRNA.
ORGANISM
Gentiana triflora
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
1 (sites)
Yonekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M.,
Fujiwara,H., Fukui,Y., Toshihiko,A., Yamaguchi,M. and Kusumi,T.
Molecular cloning and biochemical characterization of
hydroxycinnamoyl-CoA:anthocyanin
3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla
frutescens and diverse plant acyltransferase homologs
Unpublished (1999)
2 (bases 1 to 1622)
Tanaka,Y. and Yonekura-Sakakibara,K.
Direct Submission
Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Tanaka, Suntory Ltd., Institute for Fundamental Research;
Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan
(E-mail:toshikazu_tanaka@suntory.co.jp, Tel:81-75-962-8807,
Fax:81-75-962-8262)
FEATURES
Location/Qualifiers
1..1622
/organism="Gentiana triflora"
/db_xref="taxon:55190"
35..1474
/gene="GAT106"
35..1474
/gene="GAT106"
/codon_start=1
/product="acyltransferase homolog"
/protein_id="BA93452.1"
/db_xref="GI:7415597"
/translation="MAGNSEDIKVKRCRVAPPDVAEFTVPLSFDDMRWLISDAEH
HLHFYRPHPCPNKFTIISIKSLSLVLKHLPLAGNLINPVDSSORMPLRYKKGD
SVSLTIAESSMDFYLAGDHODSYKENDLIPLPELIVTSGDEVLPFLALQVFSN
TCIGIRNLHQLVGDASSFLHFNKLWLVDKNSGDSLKFLPLSLLPWDRSVQDPEH
IRRIYNERKLLKSQGTPTVLNPAISKDEVRATFLHPIDIMKLKFISSKNRLTGS
SNYLSFTVTLSALITCLSLDVTVVREKVEDKHAANLKDAIHKRLHDYEGILRGDWISPPRSTSA
APRSTLIYVYVGSQNRVHDFDADFQWGLKLEKHESVSTNPSTATLILISRSRRFKGAL
GISLPKNRMDAFATIFTNFINSLHVRPL"
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN
Query Match 9.2%; Score 138.2; DB 12; Length 1622;
Best Local Similarity 48.6%; Pred. No. 2e-20;
Matches 676; Conservative 0; Mismatches 643; Indels 72; Gaps 8;
Qy 1 TGACATTTCTCGACATCGCCGCAATATCGGCCCTCGGGCACCATCGGCATCGCTCGT 60
Db 57 TCAAAGTTCTTTGAGAAATGCGGTGTGGCCACCACCACCGCGCGTTCGCCGAGTTTACAG 116
Qy 61 TATCTCTACTTTCTTCGACATTTACTTGGCTACTTCTCCCTCCGCTCCACCATCTTTTCT 120
Db 117 TCCCACTGCTGTTTTTTCGACATGCGATGGTTGATCTCTGATGCGAGAACACCATCTGCAAT 176
Qy 121 TCTATGACTTTTCCACATCTCTAAATCCCACTTTTCATGGACACTATTGTTCCCGAGGTAAAC 180
Db 177 TCTACAGATTCGCGCATCTTGTCCCACTCTAAATTTATCATTTTCATCCA---TTAAT 233
Qy 181 AATCTTATCGGTCACTCTTCAACATTTTTTCCCGTTGGCTAGTAATTTGATTTGATTTTC 240
Db 234 CGTCCCTTTCCCTTTGTTCTCAACACACTTTCTTCGTTAGCGGGAAATTTGATTTGGCCGG 293
Qy 241 CTACACTGATGGTTCGGGTTTTTAATAAAACCAGAAATAAACACGCTTCAAGGTGATT 300
Db 294 TAGATTCCTCCG-----ATAGATGCCGGAGTTCGGTTACAAGAAAGGGACT 341

QY 301 CTGTTGTTGTTACTTTTGCAGAAATGTTGCTTGCACCTTTAATAATTTGCAGGAAATCATC 360
DB 342 CCCTGTTCTTTAAACAATTCAGAAATGCAGATGATGATTTGATTTATCTCGCGGAGATCATC 401
QY 361 CTCGAAATGTTGAAACACTTTTATCCACTTGTACCTTCATTTGGGAAATGCAATCAATTTAT 420
DB 402 AGAGGATGTTCTTATAAATTTCAAGGATTTGATTTCCGAGCTGCCAGAACCGATTTGTAACTT 461
QY 421 GTGATTTGCGTACGGTCCCACTTTTCTCACTTCAAGTGAGCTTTTTCGGGCGTCCGGTA 480
DB 462 CCGCGCAGCAAGTATTACCACTTTTTCCTTTACAGGTGACGGTGTCTCTCAACACCGGTA 521
QY 481 TATCACTAGGAATGAGCAATCATCATAGCTTGTGTGACGCTAGCAGCGGTTCAACTTTT 540
DB 522 TATGCAATGAGCACTTCTTCAAGTCTTCTGATGATGCACTTTCTTTCTGCAITTTA 581
QY 541 TGAAGGGTGGACTTCGATTTCAATCTGTTGATGATGATTTTATAAACCTTCGAGCCTTTG 582
DB 582 ATAAATTTATGGTTTGGTTGACAAATCCAAATGGAGATTCATTAAGTTCTCTTCCACTTT 641
QY 583 CTTTTTAAAGAAAGGATCTCCACCGTCTTTTGCATAGATTGATTAACATCCCACTTTAG 642
DB 642 CTTCTCTACCTATGTAGGACAGATCTGTGTGCAAGATCCATTTCAATTCGTGCAAAAA 701
QY 643 ATGAAATTAAGTTGAGACATACAGGCTCGAAAGTTTATATAACCTTCGAGCCTTTG 702
DB 702 TCTACAATGAAGAAATGCTCAAAATCTCAGGCGCACCTACTGTTCTAAATCCAGCA 761
QY 703 GTCCCACTGATAAGTTTCGTTGCTTGTGACCCGGAATCAATCAATCTACTATA 762
DB 762 TTTCTAAAGATGAAGTTTCGAGCAGCTTCTCTACACCTATTTGATATCATGAAGTCA 821
QY 763 AGAAAAGGCTTTAAACCAAGTCCCAAACTT-----GGAGTACATGTCAT 807
DB 822 AGAAATTCATTTCTCGTCAAAAATTCGCAACTTAACCGGTAGTAGTAATATATCTGTCAA 881
QY 808 CTTTTACGGTAACTTGTGTTATATATGAGTTGTCATAGCGAAATCACTGCTGAAAAATAG 867
DB 882 CTTTTCAGGTTGACATCTGCACCTGATCTGGACATGCTTGTGAAATCATATTAGACACGCTG 941
QY 868 GAGAAAGAGGGCGAAGAGC-----AGTTAGAACAGTTTCATATCAACCATTTGTTGTC 921
DB 942 TAAGAGAGAGAGTGGAGAGGATTAACATGCAGCAAACTTATGTGCTTTTCATCAACTGCC 1001
QY 922 GATCTGCTCTGTATCCACCAATTCACACGCTTCTTGTGTAAGTCTGCTGACACCTGTCG 981
DB 1002 GACAACTTTTGTCTCGCCCTACCTCAAAATTTACTTTGGAAATTCATAGTGCCTTTGTA 1061
QY 982 TCCCGACCTTAAAAAATGTCGTTTGTGACTACGGAATGGTATGCACCTTGGTCTAAAG 1041
DB 1062 TGCTGGAGTCACTCATGACCACTTGTAGGAAATGAGGTTCTCGGTAGCTGCAACCG 1121
QY 1042 TAATTGAGAGTCTATATGCAAAATGATATATATAAGGACGGAATCTTGAAGATGCCG 1101
DB 1122 CCATCGGAGATGCTATCCATAAGAGGTTTACATCACTACGAAGCAATTTCTCAGAGGAGATT 1181
QY 1102 CGAGATGGCATGAACCTTTTCATATCCCG-----CTAGGAAGATTTGGTTGCTG 1152
DB 1182 GGATATCGCCCGCGCATCAACATCTCGCGCACCAAGGTGCAGCGCTCATTTATGTCGTTG 1241
QY 1153 GTACACTTAAGCTCAACTTCTAGCACTTT-----GATTTGGTGGGGAAGCGCATAA 1206
DB 1242 GATCCGACACAGCAATGTCATGATTTTGTATGCAGATTTTGGTGGGAAACCTTGAAA 1301
QY 1207 AGTAGAGACTTTTCAATAGA---CTATAATACGCGATTTCTATAAATGCAAGCAAAA 1263
DB 1302 AGCATGAATCTGTTCAACTAATCTTTCGGCAACACATAATTTTGTCTCTCGGTCCAGAA 1361
QY 1264 CATCAGCACAGATCTTGAATTTGGATTTAGTCTACCGAGATGCAAAATGGAGCGTTTT 1323
DB 1362 GATTTAAAGGAGCACTTTGAGCTTTGGCATTTCTTTGGCTTAAGAATAGGATGGAGCATTTG 1421
QY 1324 CTAGCATCTTT 1334

DB 1422 CCACCATTTT 1432

RESULT 8
AC002560
LOCUS
DEFINITION
AC002560 101176 bp DNA PLN 19-JUL-2000
Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome 1, complete sequence.
AC002560
AC002560.2 GI:9211167
HTG.
SOURCE
thale cress.
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE
1 (bases 1 to 101176)
Khan,S., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Kim,C., Shinn,P., Altafi,H., Bei,Q., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzales,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,K., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Federspiel,N.A., Theologis,A. and Ecker,J.R.
Genomic sequence for Arabidopsis thaliana BAC F21B7 from chromosome 1
TITLE
Unpublished
2 (bases 1 to 101176)
Ecker,J.R.
Direct Submission
Submitted (25-SEP-1997) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE
3 (bases 1 to 101176)
Ecker,J.R.
Direct Submission
Submitted (13-NOV-1997) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE
4 (bases 1 to 101176)
Ecker,J.R.
Direct Submission
Submitted (26-JAN-1998) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE
5 (bases 1 to 101176)
Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
Direct Submission
Submitted (28-JUN-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th and Hamilton Walk, Philadelphia, PA 19104-6018, USA
REFERENCE
6 (bases 1 to 101176)
Ecker,J.R.
Direct Submission
Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center, Department of Biology, University of Pennsylvania, 38th Street and Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA
REFERENCE
7 (bases 1 to 101176)
Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q., Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bei,B., Chin,C., Chlou,J., Choi,E., Conn,L., Conway,A., Gonzalez,A., Hansen,N., Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J., Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H., Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M., Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.
Direct Submission
TITLE

QY 123 TATGACTTTCCACATTTCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGCTAAACAA 182
Db 37391 TTCTCTTTACCAAACTCAACTGAATCTTTCCCTCCCAAGACTTTGTACCCCAACTCAACAT 37450
QY 183 TCTTTATCGGTTCACCTCTCAACATTTTTCCTCGGTTTGGTAGTAATTTGATTTGATTTCT 242
Db 37451 TCCCTCTCCATCACTCTCCAACTTTCTTTCCCTTACGCGGGTAAATGATTTATCCCGCT 37510
QY 243 AACACTGATGGTTTGGGTTTAAATAAAACACAGAAATAAAACACGTTTCAAGGTGATTC 302
Db 37511 CGTCCGACCTCCATATTTACACTACAACGCTG---GGAAGACTCTCTTTGTTTCACC 37567
QY 303 GTTGTGGTTTACCTTTGTCAGAAATGTTGTCTGTGACTTTTAATAATTTGACAGAAATCATCT 362
Db 37568 GTAGCAGAGCTTACTACAAACCGACTTCGACCACTGATCTGATTTCCCTTAAAGATATC 37627
QY 363 CGAAATGTGAACACTTTTATCCACTTGTACCTTTCATTTGGGAATGCAATCAATATGCT 422
Db 37628 AGTGTGTTGCATGGGGTCTTGGCCCAAGTTTACCTCTCCTCACGTCTCTCCGGAAGGAAT 37687
QY 423 GATTCGGTCACGCTCCCACTTTTTCACCTTCAAGTGACGCTTTTTCGGGCTCGGCTATA 482
Db 37688 CA-----CATGGACCAATATGGAATGCAAGTCAACCATCTTCCCTGGAGCTGGAATC 37741
QY 483 TCACTAGGAATGACGAATCATCATAGCCTTGGTACGCTAGCAGCGGTTCAACTTTTGG 542
Db 37742 TGTATAGGCACTCAGCTACACATGTTGTAGCAGATGGAGTCACTTCACTCATTTTCATG 37801
QY 543 AAAGGTGGACTTCGATTAATCAATCTGTGTAGATGCGTCTTTTAAACGAAAGGATCT 602
Db 37802 AAGTATTGATGTCAATGACCAATCCAGCGCAAGATCCGCGCACGCTTCTTCTACCC 37861
QY 603 CCACCGGTTTTGCATAGATTGATTACATCC-----CACATTTAGATGAAT 650
Db 37862 TCTCTGCCATTACAGACTGCGAACAATCAATCAAGGACCCAGGAGAGTAGCCCGCAGGA 37921
QY 651 AAGTTGAGACATCAAGGCTCGAAAGTCTTTTATAAACCTTCGAGCCTTGTGTGCCACT 710
Db 37922 CATTTGGACGGTTTGGAGCCCAAACTCTGCNAACACAGTTCACATGTTAGCCCTGAG 37981
QY 711 GATAAAGTTCGGTCAACGTTTGTGTGGACCCGGAACCTAATACTACTAAGAAAG 770
Db 37982 AACATGGTCAGAGCCACTTTTACATTTAGCCGGAAGCAGATAGACAATCTCAAGAGTTGG 38041
QY 771 GTCTTAACCAAGTCCCAACTTGGAGTACATCTCATCTTTTACGCTACTTGTGGTTAT 830
Db 38042 GTTACAGACGCTGTGAGATCAATCTCTGTTTCTTACCTTCGTGTGACTCTAGCTTTC 38101
QY 831 ATATGGAGTTGCATAGCGAAATCACTCGTAAATAATAGGAGAAAGGCGGCAAGACGAG 890
Db 38102 ATTTGGGTAAGCTTGATTAGACACTTGTTCAGACAGTGAACAGAGCCCAAGGACGAA 38161
QY 891 TTAGAAGCTTCAATATCACTACCATTTGATTCGATCTCGTCTTGAT-----CCACCAATT 944
Db 38162 GTCTTTCACCTTGATGATCAACCTGGATTTCGAGGAATCGTCTCAAGTACACACCAACCA 38221
QY 945 CCCACAGCTACTTTGGTAACCTGTGTGACCATGTGTCCCGACCTTTAAATAATGCTGTT 1004
Db 38222 CCACAACATACATTTGGCAACCTGTATGGCTCTCTGTTATCGTATCATGCTCAAGAAACAG 38281
QY 1005 TTGACTACGGAATAATGGGTATGCATTTGTGTCTAAAGTAATTTGGAGGCTCTATATCAAA 1064
Db 38282 TTGCTAGGAGAAAATGCGCTATGCGTCTTTCAGATGCAATCCAGGAGATCAAGAT 38341
QY 1065 ATGATATATAAAGGACCGGAATCTTTGAAGATGCGCGGAGATGGCATGAACCTTTTCATG 1124
Db 38342 ATGTTATCAAGCATCTGTTTGAAGACAGCACCAAGATGGGACAGGAGTACGTAATGCG 38401
QY 1125 ATCCCGGCTAGGAG---ATTGGTGTGCTGGTACACCTTAAGCTCAACTTGTACGACTTT 1181
Db 38402 GTCATGCTCCATTTATCCAACTCAATTTGAGGAGCTCCGAAATTTGGGACTGTATGACATG 38461
QY 1182 GATTTTGGGTGGGGAAGCGCATAAAGTATGAGACTGTTTCAATAGACTATAATACGTCG 1241

Db 38462 GATTTTGGATTAGGAAAGCCCTGCAAAATGGAGATAGTCACATAGAAAAGGTGTTCT 38521
QY 1242 ATTTCTATAATGCAAGCAAAACATCAGCACAGATCTTTGAAATTCGATTGAGTCTACCG 1301
Db 38522 ATCGCATCTCAGAGTTCAGAGACGCGCAATGGAGTTGGAGATTGGAATAGCAGCTAGAG 38581
QY 1302 AGTATGCAATGGAGGCTTT 1322
Db 38582 AAGAAGAAATGATGTT 38602
AC003027 119914 bp DNA PLN 30-JAN-1999
Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,
complete sequence.
AC003027
AC003027.1 GI:4079614
HTG.
SOURCE
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 119914)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Unpublished
2 (bases 1 to 119914)
Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,
Oji.O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vysotskaia,V.,
Yu.G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
3 (bases 1 to 119914)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (30-DEC-1998) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
4 (bases 1 to 119914)
Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altafi,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
Direct Submission
Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
On Dec 30, 1998 this sequence version replaced gi:2734094.
Bases 1-9462 of clone F21M11 overlap with bases 68998-78259 of
'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone
F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7
(AC002560).
e-mail for correspondence: arab@sequence.stanford.edu
Genes with similarity to proteins in the databases are described as
'putative', '-like' or 'similar to'. Genes that have EST
similarity but no significant protein similarity are described as
'unknown proteins'. Genes that are annotated based only on gene
prediction software are described as 'hypothetical proteins'.
The software programs used to predict genes include: Grail

(Informatics Group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge,
<http://genome.stanford.edu/~chris/GENSCANW.html>), Fexa (V.Solovyev
 & A.Salamov, Sanger Centre, <http://genome.sanger.ac.uk/>), and
 NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, <http://www.cbs.dtu.dk/NetPlantGene.html>).

FEATURES

```

source
    location/Qualifiers
    1. .119914
    /organism="Arabidopsis thaliana"
    /cultivar="Columbia"
    /db_xref="taxon:3702"
    /chromosome="1"
    /clone="F21M11"
    1. .7000
    /note="Proteins in this region are annotated in the F21M11
    entry, AC002411."
    1. .9262
    /note="Overlap with bases 68998-78259 of 'IGF' BAC clone
    F21M11."
    7684. .10913
    /gene="F21M11.1"
    join(7684. .7866,7965. .8027,8122. .8160,8250. .8381,
    8458. .8728,8875. .9893,10003. .10256,10344. .10410,
    10501. .10551,10685. .10781,10885. .10913)
    /gene="F21M11.1"
    /note="Hypothetical protein"
    /codon_start=1
    /protein_id="AADI0665.1"
    /db_xref="GI:4204284"
    /translation="MAPNLRIKKACDAMKLLGSETKTRAPLRLKLLKTYENNNDFTBEE
    DAYKVLDAIFDEADQSTBKNKEEKKKRSVATSGRRKRAPEPLVQDDE
    DMDDEFFLRRLRRGRASSSSSSSSNNEDLQTPDEEDDQGVTELPPLKRY
    VRRNGEPLMTVYNNAPSSSRSLRMEPEVPMVLLPAHMETKYSEASALVILND
    EPNIDHKPLVTDGNCAPMLEGKSNHVGQEDWETKDLNDTAMDVSPSALICES
    SEHKVAASVELASSTGEAKICLSFAPATGETNLHLPMSDELRRAMEKCLSKSYKI
    VHPFSLVGMKMCSDITDAGNEVETVCDMSKAGDSGAGVISMPLVVVP
    ECEISGDGKAISNDITAGEENVEIPWNEINEKPSRFYMPHSFVQDAPVIFS
    LSSFQDSQSTCIEDCLASGNCAGIDNGFAYTLDLGLKEEFLEARISEARDQ
    RKQVRECEPCEPRAKKVLEIPCKGLKRGAKICEKFCGCTKCGNRVVGORMHN
    KLOVFEPNGKGLRTELEKLPKGAFCIEVIGELTIPELYQSFEDKPTLPVILDAH
    WGESELEGDKALCDGMFYGNISFLNRLCLDANLIEIPQVETPDQHYHLAFPTT
    RDPEAMELAWGIDRNDMSLMKPFDCIGSGFRCKRRKRTKMQLNKA"
    join(13033. .13488,13717. .13926,14139. .14514)
    /gene="F21M11.2"
    13033. .14514
    /gene="F21M11.2"
    join(13061. .13488,13717. .13926,14139. .14316)
    /gene="F21M11.2"
    /note="Similar to acid phosphatase; Location of ESTs
    110C2T7 , gb|T42036, and 110C2XP, gb|A1100245"
    /codon_start=1
    /protein_id="AADI0666.1"
    /db_xref="GI:4204285"
    /translation="MDRTMFLSTIALLGVVVSAGDWNILNOLRGLGSSSSONGIVS
    KGRKDLKGCESNRINVEYNNIRKEDVPQECVSHIKDYMTSQYKDDVARTVDEVI
    LHFSGMCCSKCDGMDAIFDIDTLLSTIPYHKKGFFGKLNSTKFEEDWTQKK
    APVAPHMKLVHLDIRERGIKELISSRKEYLSATVDNLQAGYIGNSNMLRGLDEQ
    QKEVKQYKSEKRWLMSLGYRVWVGMDGDSRAGCPLPRTPKLPNSIYVA"
    complement(join(14004. .14312,14565. .14641,14787. .15831,
    15906. .16006,16327. .16362))
    /gene="F21M11.3"
    15906. .16006,16327. .16362)
    /gene="F21M11.3"
    complement(14004. .16362)
    /gene="F21M11.3"
    complement(join(14567. .14641,14787. .15831,15906. .16006,
    16327. .16362))
    /gene="F21M11.3"
    /note="Unknown protein; Location of ESTs 40C3TT,
    gb|AA728590 and 40C3T7, gb|T04573"
    /codon_start=1
    /protein_id="AADI0667.1"
    /db_xref="GI:4204286"
    /translation="MLDQISGNHTHEKLSVETAEPHLNDNRVHIVEIPKASVITPE
    ICDAEKEKSPSTISRKRVTDSKVKTYSYHVWSEESVELSEKNEEVEKRLSKSK
  
```

gene

CDS

```

    complement(join(16835. .17185,17274. .17392,17491. .17656,
    17795. .17885,17982. .18079,18175. .18361,18504. .18604,
    18705. .19049,19134. .19349,19439. .19791,19862. .19967,
    20188. .20238))
    /gene="F21M11.4"
    /note="Hypothetical protein"
    /codon_start=1
    /protein_id="AADI0669.1"
    /db_xref="GI:4204288"
    /translation="MGSSSPPEARARAOVPSMILIFLEIITVHYVYNNRRKLNDRDLVLSA
    NLNIPKRVTRDLRFMESVSIYKSLNAAVSVHLPOQIPNDCNYKNDALNNSNPKH
    GESEDSMTDKVSKRSGTSSSRDGLPTSESDPRKHQDQTEKQLSHLLYE
    FESEYDAANHTPESYTBQAAKNVDRDITASEQPSNAARKRICGDSFIOESSNPKTQDP
    TLLRLLESLSDDOPTDYKQONHQMVYSEFWLYSFVLCILRSNNLRVCWIFCVOQLPK
    SHTEQDSKRRDRITASDAMENHLKVPKRENNLMQKADIDCNKCSANSDQDSEKIS
    KALEQTSNITICGQCSARVSEATGEMLHYSRGPVDDGDIFRSNVIHVHSACIEWA
    PQVYEGDTYKNLKAEALARGMKIKCTKSLKGAALGCFVKSCRRSLEQTPAFTKELVCGR
    WQYEDFLLCPAHSSVFKPNEKSHRVSRAEPLPKINPAELCSLEQTPAFTKELVCGR
    SALSKSDKLMLESIAVRFNATISRYWNPSTVTHVIASTDEKGACTRTLKVLMLGKNGW
    IINRAMKASLKASQVDEEPELQIDTCQCGQGPRTARLAETNPKLEGLKFFYF
    GDYKTEKDELQNLVAGGTILNDELGAESSNNVNDORSISVIVINDIPPHGCAL
    GEVTTIIRANDAEALASQTSRLVGTWVLESIAGYKLHPVIG"
    20752. .24647
    /gene="F21M11.5"
    join(20752. .20994,21362. .21497,21596. .21740,21825. .21994,
    22102. .22178,22542. .22758,22920. .23064,23343. .23442,
    23599. .23693,24021. .24072,24227. .24298,24441. .24647)
    /gene="F21M11.5"
    /note="Hypothetical protein"
    /codon_start=1
    /protein_id="AADI0668.1"
    /db_xref="GI:4204287"
    /translation="MGANSKSVTASFVIAVFFLICGRTAVDETEFEHGYSKLSGI
    IIPFASTQURAWSILDCPYTDLDFNPLDLWDLTTKLLSAVNCWFKCMVLDPNQTD
    HPCKSRPDSGLSAITELDPGYITGPLSTVYKWLKWCVEFGIEANAIVAPVDMRLS
    PTKLEERDLYFHKLKLTFTALRGLGPSIVFAHSMGNVFRYLEWLRLLEIAPKHYL
    KWLDOHITHAYFAGPLLGVEAKISTLSGVTGLPVSEGTNLSFASASSLWLMFP
    SKNCKGNTFTWTFHSGCAAKKRVYHCDDEEYQSKYSGMPTNINTEIPSTARELA
    DGLFKAIEDYDPSDKRMLHQLKKYVPFFVIRNIAHRSSLAGELLYHDDDFVFNPLTPW
    ERPIKRVCIYGAHLKTEVGYIFAPSGRYPDNWIIITIIIEGSLVSRSGTVVDG
    NAGPITGDETPYVHSLSWCKNLWLGPKVNIITMAFQILGKIQQPEHSDGSDVHVELNVD
    HEHGSDIANNTRKAPRVKITYFVDESISIPGKRTAVWELDKSGY"
    27777. .28734
    /gene="F21M11.6"
    27777. .28734
    /gene="F21M11.6"
    28007. .28465
    /gene="F21M11.6"
    /note="Unknown protein; Location of ESTs 203I24T7,
    gb|H76794 and 203I24XP, gb|AA605510"
    /codon_start=1
    /protein_id="AADI0670.1"
    /db_xref="GI:4204289"
    /translation="MNTKTMRLPPRRVLTADKKRERDAFISSTVDNPPETAKFPSPPP
    KLVPPVNPISKSSATAAAGISNOLMAGYLSHEVLTOTGLFGEOGNARAQAES
    KIKPSHTVEPAEECEPKRKRYREVANLLRSDGQALPGIVNPAQLARFLKL"
    complement(29264. .32033)
    /gene="F21M11.7"
    complement(join(29264. .31015,31312. .31414,31484. .31587,
    31782. .32033))
  
```

gene

CDS

```

    complement(join(16835. .17185,17274. .17392,17491. .17656,
    17795. .17885,17982. .18079,18175. .18361,18504. .18604,
    18705. .19049,19134. .19349,19439. .19791,19862. .19967,
    20188. .20238))
    /gene="F21M11.4"
    /note="Hypothetical protein"
    /codon_start=1
    /protein_id="AADI0669.1"
    /db_xref="GI:4204288"
    /translation="MGSSSPPEARARAOVPSMILIFLEIITVHYVYNNRRKLNDRDLVLSA
    NLNIPKRVTRDLRFMESVSIYKSLNAAVSVHLPOQIPNDCNYKNDALNNSNPKH
    GESEDSMTDKVSKRSGTSSSRDGLPTSESDPRKHQDQTEKQLSHLLYE
    FESEYDAANHTPESYTBQAAKNVDRDITASEQPSNAARKRICGDSFIOESSNPKTQDP
    TLLRLLESLSDDOPTDYKQONHQMVYSEFWLYSFVLCILRSNNLRVCWIFCVOQLPK
    SHTEQDSKRRDRITASDAMENHLKVPKRENNLMQKADIDCNKCSANSDQDSEKIS
    KALEQTSNITICGQCSARVSEATGEMLHYSRGPVDDGDIFRSNVIHVHSACIEWA
    PQVYEGDTYKNLKAEALARGMKIKCTKSLKGAALGCFVKSCRRSLEQTPAFTKELVCGR
    WQYEDFLLCPAHSSVFKPNEKSHRVSRAEPLPKINPAELCSLEQTPAFTKELVCGR
    SALSKSDKLMLESIAVRFNATISRYWNPSTVTHVIASTDEKGACTRTLKVLMLGKNGW
    IINRAMKASLKASQVDEEPELQIDTCQCGQGPRTARLAETNPKLEGLKFFYF
    GDYKTEKDELQNLVAGGTILNDELGAESSNNVNDORSISVIVINDIPPHGCAL
    GEVTTIIRANDAEALASQTSRLVGTWVLESIAGYKLHPVIG"
    20752. .24647
    /gene="F21M11.5"
    join(20752. .20994,21362. .21497,21596. .21740,21825. .21994,
    22102. .22178,22542. .22758,22920. .23064,23343. .23442,
    23599. .23693,24021. .24072,24227. .24298,24441. .24647)
    /gene="F21M11.5"
    /note="Hypothetical protein"
    /codon_start=1
    /protein_id="AADI0668.1"
    /db_xref="GI:4204287"
    /translation="MGANSKSVTASFVIAVFFLICGRTAVDETEFEHGYSKLSGI
    IIPFASTQURAWSILDCPYTDLDFNPLDLWDLTTKLLSAVNCWFKCMVLDPNQTD
    HPCKSRPDSGLSAITELDPGYITGPLSTVYKWLKWCVEFGIEANAIVAPVDMRLS
    PTKLEERDLYFHKLKLTFTALRGLGPSIVFAHSMGNVFRYLEWLRLLEIAPKHYL
    KWLDOHITHAYFAGPLLGVEAKISTLSGVTGLPVSEGTNLSFASASSLWLMFP
    SKNCKGNTFTWTFHSGCAAKKRVYHCDDEEYQSKYSGMPTNINTEIPSTARELA
    DGLFKAIEDYDPSDKRMLHQLKKYVPFFVIRNIAHRSSLAGELLYHDDDFVFNPLTPW
    ERPIKRVCIYGAHLKTEVGYIFAPSGRYPDNWIIITIIIEGSLVSRSGTVVDG
    NAGPITGDETPYVHSLSWCKNLWLGPKVNIITMAFQILGKIQQPEHSDGSDVHVELNVD
    HEHGSDIANNTRKAPRVKITYFVDESISIPGKRTAVWELDKSGY"
    27777. .28734
    /gene="F21M11.6"
    27777. .28734
    /gene="F21M11.6"
    28007. .28465
    /gene="F21M11.6"
    /note="Unknown protein; Location of ESTs 203I24T7,
    gb|H76794 and 203I24XP, gb|AA605510"
    /codon_start=1
    /protein_id="AADI0670.1"
    /db_xref="GI:4204289"
    /translation="MNTKTMRLPPRRVLTADKKRERDAFISSTVDNPPETAKFPSPPP
    KLVPPVNPISKSSATAAAGISNOLMAGYLSHEVLTOTGLFGEOGNARAQAES
    KIKPSHTVEPAEECEPKRKRYREVANLLRSDGQALPGIVNPAQLARFLKL"
    complement(29264. .32033)
    /gene="F21M11.7"
    complement(join(29264. .31015,31312. .31414,31484. .31587,
    31782. .32033))
  
```

Query Match

5.88; Score 88.2; DB 12; Length 119914;

Best Local Similarity 45.08; Pred. No. 1.9e-09;

Matches 621; Conservative 0; Mismatches 718; Indels 42; Gaps 6;

Db	1002	CAAGCCAAAACCTGAAGCTGAAGCTGCAGCAGGTAATGATGAAATTAATAATATTATTGT	1061
Qy	902	CATAATCACCATTGATTGTCGATCTCGTCTGTGATCCACCAATTCACACAGCCTACTTTGG	961
Db	1062	GGGATTCACTCGGATTATAGGACCGCTTTGGATCCTCCAAATTCACACTTAATTATTTGG	1121
Qy	962	TAACGTGGTGACCATGTGTCGCCAGCTTAAATAATGTCGTTTGTAGCTACGGAATAATGG	1021
Db	1122	TAACGTGAATGGAGACATTTGTGAGACTGCAAAAGCAAGTGATTTCTGTTCAAGAAAATGG	1181
Qy	1022	GTATGCACCTGGTCTAAAGTAATTTGGAGA--GTCATATATGCAAAATGATATAATAA	1078
Db	1182	GTTGCTCTTTGTCGAGAGATGTTAAAGTGATATGGTCAAGGGATTCGATGCGGATGCCAT	1241
Qy	1079	GGACGAATCTTGAAGATGCCGAGATGGCATGAACCTTTTCATGATCCCGGCTAGGAA	1138
Db	1242	TGAAGCCAAATGATGATAAGCTTTTCAGAAATATTGGAATTTCTGAAGAAGGACCAATGAT	1301
Qy	1139	GATTGGTGTGCTGGTACACCTAAGCTCAACTGTTACGACTTTTGATTTGGTGGGGGAA	1198
Db	1302	TTTTTCTGCTCGCTCGACCCAAATTTGATGTTTACGGTTCGGATTTTCGGTGGGGAG	1361
Qy	1199	GGCATAAAGTATGAGACTGTTTCAATAGACTATAATACGTCGATTTCTATAATGCAAG	1258
Db	1362	GGCCAAAGAGTGGAGATGTTGTCATATAGTAGGACACAAGCCATCTCTTTGGCAGAGAG	1421
Qy	1259	CAAAACATCAGCACAAAGATCTTGAATTTGGATTGAGTCTACCGAGTATGCAAAATGGAGGC	1318
Db	1422	AAGATGGAGGAGCGCGCTTGAGTTGGAGTTGTTTATAGAGAACAACAATTTGAGGT	1481
Qy	1319	GTTTCTAGCATCTTTGATGAAGGATTAGAGAGTCAAGTTTTCATTTGATGATCAT	1372
Db	1482	TTTTGAATCTGATTGCTGATGGACTGAAATAATGATCTCTGTTTAATTAATGAT	1535
RESULT 11			
AB016892			
LOCUS	AB016892 66237 bp DNA PLN 27-DEC-2000		
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MXF12.		
ACCESSION	AB016892 BA000015		
VERSION	AB016892.1 GI:3449333		
KEYWORDS			
SOURCE	Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1 clone:MXF12.		
ORGANISM	Arabidopsis thaliana		
	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.		
REFERENCE	1 (sites)		
AUTHORS	Asamizu,E., Sato,S., Kaneko,T., Nakamura,Y., Kotani,H., Miyajima,N. and Tabata,S.		
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. VIII. Sequence features of the regions of 1,081,958 bp covered by seventeen physically assigned P1 and TAC clones		
JOURNAL	DNA Res. 5 (6), 379-391 (1998)		
MEDLINE	99156233		
REFERENCE	2 (bases 1 to 66237)		
AUTHORS	Nakamura,Y.		
TITLE	Direct Submission		
JOURNAL	Submitted (18-AUG-1998) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp, Tel:81-438-52-3935, Fax:81-438-52-3934)		
	Address for correspondence: kaos@kazusa.or.jp		
COMMENT	For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/agd-graph.cgi?c=MXF12		
	Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'.		
	The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Grail-1.3/).		
	GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),		

NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlin.cool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K15E6 and the 3' clone is K3K3.	
FEATURES	Location/Qualifiers
source	1..86237
	/organism="Arabidopsis thaliana"
	/strain="Columbia"
	/db_xref="taxon:3702"
	/chromosome="5"
	/clone="MXF12"
	/clone_lib="Mitsui P1"
CDS	2613..5234
	/note="gene_id:MXF12.2"
	/codon_start=1
	/evidence=not_experimental
	/product="receptor-like protein kinase"
	/protein_id="BAB10824.1"
	/db_xref="GI:10177545"
	/translation="MIRHALLIFSLVSTPIVGBGATSTVEPTDVFLENGCDTSNNVDVSRNNTAENQKILSSNLVNAFTSAQSYOESGVSQIPYMTARIFRSEFTYSPPVTPGSGNELRLFYPTRYGSPQNAVKSFFSVKVNGETLLNFSADLTVKASKPQTEFIKEFIIPVQTLNLTFTPSDSLAFVNGIEIVSNRFSYKGGFDDVTINVGSSVDFHLENSTAFETVYRLNKGKTVGSDGMRFRWVSDEILISSESGISIPVDIKINTEKTPSYVAPDDVIATSRMGNAHPENLNLWLTFTVDAGFSYLVRHLRDLHLPLNVSINPKYSFIENOTATLEMDVFRMSGSWIPMYLDYTVIAGSGRRHDLRLDHLPLNVSINPKYVAILNGVEILKMNPDGNLAGNPDLVSPDLIPNRTAIPRKNSHILPITLAVVGSFLVLANFVGVGLVIMKKKKSPSTNSNCPDPHGTDSTNTPAKSLPADLCRRFSIFEIKSATNDFEDKLIIGGFGSVYKQIDGGATLVAKREITISNGAKEFELSEM LSKLRHVLHSLIGYCEDENMLVTEYMPHGTLDHLFRDKTSDPLSKRRLEICIGARGLOYLHTGAKYTIHRDITKLTILLDENFTVSDFGLSRVGPTSAQTHVSTVNGTFCGLDPEYRROVLTEKSDVSFGVVLLEVLCCRPTRMOSVPEQADLIRWVKSNNERTGVDOIISDLSADJTSLEKECIAVRCVQDRGMERPMDNDVYVALEFALQLHETAKKKNONVESLDLMPSEGEVGTITTDGDDLLFSRTTGHVGKSTTTDDSLVVGDERSGSSWGVFSEINEPKAR"
CDS	6701..7210
	/note="gene_id:MXF12.3
	unknown protein"
	/codon_start=1
	/evidence=not_experimental
	/protein_id="BAB10825.1"
	/db_xref="GI:10177546"
	/translation="MINLEKAKEKQILIRDKRKIVHHVNTTLLVKIYEASAIIDRRHLFLKPYGPPNYVFIKAWLDDTKGYTVYVVKDSFKVNOPPLKIPLNFPADQVLYVELLK GMSYRDPTGTSNRTVVMGRAKIRLSPLMTSYKERPISGGIDLVGLNSDQCVVKGGLTYAKMLHRSAD"
CDS	7670..10111
	/note="gene_id:MXF12.4"
	/codon_start=1
	/evidence=not_experimental
	/product="receptor protein kinase-like protein"
	/protein_id="BAB10826.1"
	/db_xref="GI:10177547"
	/translation="MNCNVLFLSVLSVGTAGVTAAVHPTDVFLENGCDTSNNVDSNGRNTVRESQILSSNLVNAFTSEASYQAGVSRIPYMKARIFRSEFTYSPPVTPGSGIFLRLYFTPTQYKSGFDVANSFVSXVNGFTLLRNFNADSTVQASIPLSNLSIKKEFIIPVHOTLNLTFTPSKNLLAFVNGIEIVSNRFSYKGGFDDVLRNVSDDVDFDIDNSTAFESVHRLNNGGOIVNEVDSDGMRFRWLSDDSEFGSGSIVNPGVKINTEKTPYAVPDVMATSRMGNSNLMFNLTGCMFLITVDAGYNLVRHLCFCLPQVTKAGQVRSFIIVDDKMAKSTDVIRLSSGPRIPYLDFTSVTVYFESGMIOPELRLDVLKPDQVNTQYTDAILSGVEILKNDSDGNLARNPNPELLVSDTSDPDSNVPTPIKGRPHVLVILTVGSGVIGLATFIVTILLIRQMKRKNKNSITMFKLLKQYTYAELKTKTSFSTVKGSGVGGGTVYVRGNLSNGRTVAVKVLKDKNGDDFINVETSMSTSHVNIYSLGLFCVYEGSKRAI ISEFLEHSGLDQFISLRNKSITPNVTTYLGALGIAGLEVLYHVCCKTRIVHFDIKPON ILDDNFCPKVADFGALKCEKRESILSLIDTRGTIGYIAPEVVSRYMGGISHSKSDVISYGMVLDMIGARKNKVEFTTTPDWIYKOLENGDOTWIIGDEINEEDNKKIVK KMLVLSLWCIRPCPSDRPPMNKVMETEGSLDALELPKPPSPRHSIISTELVLESSLSDC

NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, <http://www.cbs.dtu.dk/services/NetGene2/>) and SplicePredictor (Volker Brendel, Stanford University, <http://gremlin1.zool.tastate.edu/cgi-bin/sp.cgi>) Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K1566 and the 3' clone is K3K3.

Db 351 ATWWWTTTAKWGTATATKWTAKTKTTAKTKTTTWTGTTTWTGTTTWTGTTTAKTKTKTKTTTW 292

Qy 1411 ATGAGTGGCAGCTGCTTCTATTCTTTAAGTATACCTTTCGACTATGTTTGAAGATGCAAC 1470

Db 291 AWTTTGKTATATKWTATKATGAGGAKWTTTGGKACATTTTTTTTGGTGGKAWTKT 232

Qy 1471 GATATAAATGAAAAAAMAAAAAAMAAAAAAMAAAAA 1508

Db 231 WWWTTTAAAAAAWATATAAAAAAAMAAAAAAMAAAAA 194

RESULT 14

LOCUS I66494 7218 bp DNA PAT 28-DEC-1997

DEFINITION Sequence 14 from patent US 5670367.

ACCESSION I66494

VERSION I66494.1 GI:2724471

KEYWORDS

SOURCE Unknown.

ORGANISM Unknown.

REFERENCE 1 (bases 1 to 7218)

AUTHORS Dörner,F., Scheiflinger,F. and Falkner,F.Gunter.

TITLE Recombinant fowlpox virus

JOURNAL Patent: US 5670367-A 14 23-SEP-1997;

FEATURES

Location/Qualifiers

Source 1..7218

/organism="unknown"

BASE COUNT 1944 a 1491 c 1486 g 1929 t 368 others

ORIGIN

Query Match 3.5%; Score 53.6; DB 10; Length 7218;

Best Local Similarity 1.9%; Pred. No. 0.083;

Matches 5; Conservative 171; Mismatches 90; Indels 0; Gaps 0;

Qy 7 TTCTCGAACATGCCGAATATCGGCCCTCGCGCCACATCGCCATCGCTTATCTC 66

Db 1177 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1236

Qy 67 TACTTTCTTCGACATTACTTGCTACTCTCCCTCGGTCGCCACCATCTTTCTCTATG 126

Db 1237 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1296

Qy 127 ACTTTCACATCTTAATCCATTTTCATGGACACTATTGTTCCAGGCTAAACAATCTT 186

Db 1297 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1356

Qy 187 TATCGGTCACTCTCAACATTTTTCCTCGTTTGTCTAGTAATTTGATGTATTCCTAACA 246

Db 1357 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1416

Qy 247 CTGATGTTCCGGTTTTAATAAAAAA 272

Db 1417 YYYYYYYYYYYYYYYYYYGTACCAAA 1442

RESULT 15

PFMAL4P3

LOCUS

DEFINITION Plasmodium falciparum chromosome 4 strain 3D7, *** SEQUENCING IN

PROGRESS ***, in unordered pieces.

ACCESSION AL035476

VERSION AL035476.6 GI:5763808

KEYWORDS HTG; HTGS_PHASE1.

SOURCE malaria parasite P. falciparum.

ORGANISM Plasmodium falciparum

REFERENCE 1 (bases 1 to 321003)

AUTHORS Bowman,S., Churcher,C., Harris,B., Harris,D., Lawson,D., Quail,M. and Barrell,B.

TITLE Direct Submission

JOURNAL Submitted (24-FEB-1999) P.falciparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK

COMMENT On Aug 24, 1999 this sequence version replaced gi:5731899. For more information about this sequence or the Malaria Project, see http://www.sanger.ac.uk/projects/P_falciparum. IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continues. The sequence may be contaminated with foreign sequence from E.coli, yeast, vector, phage etc.

Order of segments is not known; 800 n's separate segments.

* NOTE: This is a 'working draft' sequence.

* This record will be updated with the finished sequence

* as soon as it is available and the accession number will

* be preserved.

FEATURES

Location/Qualifiers

Source 1..321003

/organism="Plasmodium falciparum"

/strain="3D7"

/db_xref="taxon:5833"

/chromosome="4"

BASE COUNT 127214 a 30738 c 31419 g 123572 t 8060 others

ORIGIN

Query Match 3.5%; Score 53.2; DB 84; Length 321003;

Best Local Similarity 49.6%; Pred. No. 0.092;

Matches 136; Conservative 0; Mismatches 138; Indels 0; Gaps 0;

Qy 1235 TAGTCGATTTCTATAAATGCAAGCAAAACATCAGCAACAGATCTTCAAATTTGGATTGAG 1294

Db 81849 TACATATATTTTCGAGAGAGGGGAAAAAGAAAAAATAATAGAAATTTATATA 81908

Qy 1295 TCTACCGAGTATGCAAAATGGAGGCGTTTTCTAGCATCTTTTGATGAAGGATTAGAGACTCA 1354

Db 81909 TATATAAGAAATGAAAAATGGAAATTTCTTTTATAATATATATGACATTATATTA 81968

Qy 1355 AGTTTCATCTAGATCATCGTCCCTTTTGTGTCGATCAAGTTTCTGCTGTTTTATGA 1414

Db 81969 AGTTTATATTATTTAATTTTCTTTTAAAAATATTATTTATTTTTTTTTTTT 82028

Qy 1415 GTTGCCACTGCTCTATTCTTTAAGTATACCTTTTCGACTATGTTTGAAGATGCAACGATA 1474

Db 82029 CTCCTATAAATAATAAG 82088

Qy 1475 TAAAAATGAAAAAAMAAAAAAMAAAAAAMAAAAA 1508

Db 82089 GTRAAGCAACAAAAAAMAAAAAAMAAAAAAMAAAAA 82122

Search completed: November 5, 2001, 16:54:14

Job time: 11024 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:47 ; Search time 445.49 seconds
(without alignments)
2125.470 Million cell updates/sec

Title: US-08-894-356C-5
Perfect score: 1508
Sequence: 1 TGAACATTCGACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_0601.*
1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1500	99.5	1508	17	AAT37312 Aromatic acyl tran
2	205.6	13.6	1703	17	AAT37308 Aromatic acyl tran
3	185.8	12.3	1479	17	AAT37311 Aromatic acyl tran
4	166.6	11.0	1518	17	AAT37313 Aromatic acyl tran
5	136.6	9.1	1622	17	AAT37309 Aromatic acyl tran
6	121.6	8.1	936	22	AAF58252 Oligonucleotide D1
7	121.6	8.1	936	22	AAF58254 Oligonucleotide D1
8	121.6	8.1	936	22	AAF58257 Oligonucleotide D1
9	121.6	8.1	936	22	AAF58259 Oligonucleotide D2
10	121.6	8.1	936	22	AAF58262 Oligonucleotide D2
11	121.6	8.1	938	22	AAF58255 Oligonucleotide D1

12	118.2	7.8	936	22	AAF58252	Oligonucleotide D1
13	118.2	7.8	936	22	AAF58254	Oligonucleotide D1
14	118.2	7.8	936	22	AAF58257	Oligonucleotide D1
15	118.2	7.8	936	22	AAF58259	Oligonucleotide D2
16	118.2	7.8	936	22	AAF58262	Oligonucleotide D1
17	118.2	7.8	938	22	AAF58255	Oligonucleotide D1
18	86	5.7	1648	21	AAF64785	Lemon acyl transe
19	53.8	3.6	1379	21	AAC46852	Arabidopsis thalia
20	52.6	3.5	244	22	AAF58238	Oligonucleotide D1
21	50.8	3.4	860	21	AAC74426	Human secreted pro
22	48.2	3.2	1469	21	AAC95491	Human secreted pro
23	48.2	3.2	1925	21	AAC78138	Human cancer assoc
24	48	3.2	244	22	AAF58238	Oligonucleotide D1
25	47	3.1	1285	21	AAC59828	Human secreted pro
26	47	3.1	1875	19	AAV33033	Dirofilaria immiti
27	47	3.1	1875	19	AAV33017	Dirofilaria immiti
28	46.6	3.1	1534	17	AAT15995	Human interleukin-
29	46.2	3.1	4391	18	AAT60666	Mouse CD100 antige
30	45.8	3.0	3184	21	AACT75894	Human ORFX ORF1449
31	45.2	3.0	1140	21	AAC98004	Human colon cancer
32	45.2	3.0	1377	22	AAF44479	Mouse dextran sodi
33	45.2	3.0	1974	21	AAAS4282	Human mesenchymal
34	45	3.0	1122	21	AACT59379	Human secreted pro
35	44.8	3.0	259	18	AAV00423	3' fragment of clo
36	44.8	3.0	1694	20	AAV07431	Homo sapiens secre
37	44.8	3.0	2907	21	AAC98917	Human pancreatic c
38	44.6	3.0	637	21	AAC79045	Human secreted pro
39	44.6	3.0	858	21	AAC59226	Human secreted pro
40	44.6	3.0	1641	17	AAT15758	Human interleukin-
41	44.4	2.9	2163	21	AAAF21924	Human breast and o
42	44.2	2.9	3495	20	AAV33814	Coding sequence fo
43	44.2	2.9	3521	14	AAQ41226	Clone GP3 encoding
44	44	2.9	1694	19	AAV33190	Secreted protein B
45	44	2.9	2497	21	AAC61197	Human cDNA encodin

ALIGNMENTS

RESULT 1
AAT37312
ID AAT37312 standard; cDNA to mRNA; 1508 BP.
AC AAT37312;
XX
DT 06-FEB-1997 (first entry)
XX
DE Aromatic acyl transferase coding sequence.
XX
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KW plants; acylation; colour: tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
KW Senecio cruentus; Lavandula angustifolia; ds.
XX
OS Senecio cruentus (Clone pCat8).
XX
FH Key Location/Qualifiers
CDS 3..1367
FT /*tag= a
FT /product= Aromatic acyl transferase.
XX
PN WO9625500-A1.
XX
PD 22-AUG-1996.
XX
PF 16-FEB-1996; 96WO-JP00348.
XX
PR 30-JAN-1996; 96JP-0046534.
PR 17-FEB-1995; 95JP-0067159.
PR 29-JUN-1995; 95JP-0196915.
XX
XX (SUNR) SUNTORY LTD.
XX

PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
 PI Nakao M, Tanaka Y, Yonekura K;
 XX WPI; 1996-393401/39.
 DR P-PSDB; AAW04726.

XX DNA coding for aromatic acyl transferase - for transforming plants
 PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers

PS Claim 4; Page 69-72; 94pp; Japanese.

XX Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313.

XX Sequence 1508 BP; 442 A; 294 C; 295 G; 477 T; 0 other;

Query Match 99.5%; Score 1500; DB 17; Length 1508;
 Best Local Similarity 99.7%; Pred. No. 0;
 Matches 1503; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 TGAACATTCTCGAACATGCCGGAATATCGGCCCTCGGCACCATCGGCCATCGCTCGT 60
 Db 1 tgaacattctcgaaacatgccgaatatcgccccctcgggcaaccatcgccatcgctcgt 60
 QY 61 TATCTCTTACTTCTCGACATTACTTGGCTACTTCCCTCCGGTCCACCATCTTTTCT 120
 Db 61 tatctcttacttctcgacattacttggctacttctccctccgggtccaccatctttct 120
 QY 121 TCTATGACTTTCACATTTCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGCTAAAC 180
 Db 121 tctatgacttctcacatTTCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGCTAAAC 180
 QY 181 AATCTTTATCGGCTACTCTTCAACATTTTTCCTGCTAGTAGTAATTTGATTTTC 240
 Db 181 aatctttatcggtactcttcaacatTTTTCCTGCTAGTAGTAATTTGATTTTC 240
 QY 241 CTAACACTGATGGTTCGGGTTTATAAATAAACACAGAAATAAACACGTTGAAGTGATT 300
 Db 241 ctaacactgatgggttcgggtttataaataAACACAGAAATAAACACGTTGAAGTGATT 300
 QY 301 CTGTTGTGGTACTTTTGCAGAAATGTTCTGTGACTTTTAATAATTGACAGAAATATC 360
 Db 301 ctgttgtggtacttttgcagaaatgTTCTGTGACTTTTAATAATTGACAGAAATATC 360
 QY 361 CTCGAAATGTGAACATTTTATCCACATTTGACCTTCAATTTGGGAATGCAATCAATPAT 420
 Db 361 ctcgaaatgtgaacatTTTATCCACATTTGACCTTCAATTTGGGAATGCAATCAATPAT 420
 QY 421 GTGATTGGCTCAGGTCCTCCACTTTTTCACCTCAAGTCAAGTCAAGTCAAGTCAAGT 480
 Db 421 gtgattggctcaggctcctccactTTTTCACCTCAAGTCAAGTCAAGTCAAGTCAAGT 480
 QY 481 TATCACTAGGAATACGAATCATATAGCTTGGTACGCTAGCACGCGGTTCAACTTTT 540
 Db 481 tatcactaggaaatacgaatCATATAGCTTGGTACGCTAGCACGCGGTTCAACTTTT 540
 QY 541 TGAAGGTTGACTTCGATTATTTCAATTCGTTGATAGTCGCTCTTTTAAACGAAAGAT 600
 Db 541 tgaagggttgacttcgattatTTTCAATTCGTTGATAGTCGCTCTTTTAAACGAAAGAT 600
 QY 601 CTCACCGGTTTTTCATAGATTGATTACATCCACATTTTATAGATGAATAAGTTGACAC 660
 Db 601 ctccacgggttttcatagattgattacacatccacatTTTATAGATGAATAAGTTGACAC 660
 QY 661 ATACAAGGCTCGAAAGTTTTTATAAACCTTCGAGGCTTTGGTGGCTCCCACTGATAAAGTTC 720
 Db 661 atacaaggctcgaaagTTTTTATAAACCTTCGAGGCTTTGGTGGCTCCCACTGATAAAGTTC 720

Db 661 atacaaggctcgaaagttttttataaaacctcgagccttctgttgggtcccaactgataaagttc 720
 QY 721 GGTCAACGTTTGTGTGACCGGAACATAATCAATCTACTAAAGAAAGAGTCTCTAAACCC 780
 Db 721 ggtcaacgTTTGTGTGACCGGAACATAATCAATCTACTAAAGAAAGAGTCTCTAAACCC 780
 QY 781 AAGTGCCTAACTTGGAGTACATGTCATCTTTTACGGTAACTTTGTTGTTATATATGAGATT 840
 Db 781 aagtgcctaaactTGGAGTACATGTCATCTTTTACGGTAACTTTGTTGTTATATATGAGATT 840
 QY 841 GCATAGCGAAATCACTCGTAAATAATAGGAGAAAGAGCGGAACGAGTTAGAACAGT 900
 Db 841 gcatagcgaaatCACTCGTAAATAATAGGAGAAAGAGCGGAACGAGTTAGAACAGT 900
 QY 901 TCATAATCACCATTGATTGTGATCTCGTCTGTATCCACCAATTTCCACAGCTACTTTTG 960
 Db 901 tcataatcaccattGATTGTGATCTCGTCTGTATCCACCAATTTCCACAGCTACTTTTG 960
 QY 961 GTAACGTGGTGCACCATGTCGCGGACCTTTAAATAATGTCGTTTGTGACTCGGAAATG 1020
 Db 961 gtaacgtggTGCACCATGTCGCGGACCTTTAAATAATGTCGTTTGTGACTCGGAAATG 1020
 QY 1021 GGTATGCACTTGGTCTAAAGTAAATTGGAGAGTCTATATGCAAAATGATATATAAAGG 1080
 Db 1021 ggtatgcaCTTGGTCTAAAGTAAATTGGAGAGTCTATATGCAAAATGATATATAAAGG 1080
 QY 1081 ACGGAATCTTGAAGATGCCGAGATGCGATGAACCTTTTCATGATCCCGCTAGGAAGA 1140
 Db 1081 acggaatctTGAAGATGCCGAGATGCGATGAACCTTTTCATGATCCCGCTAGGAAGA 1140
 QY 1141 TTGGTGTGCTGGTACACCTTAAGCTCAACTTGTAGGACTTTTGATTTGGGTGGGGAAGC 1200
 Db 1141 ttggTGTGCTGGTACACCTTAAGCTCAACTTGTAGGACTTTTGATTTGGGTGGGGAAGC 1200
 QY 1201 GCATAAGATATGAGACTGTTTCAATAGACTATAATACCTGATTTCTATAAATGCAAGCA 1260
 Db 1201 gcataaGATATGAGactgtttcaatagactataaTACCTGATTTCTATAAATGCAAGCA 1260
 QY 1261 AAACATCAGCACAAGATCTTGAAATTTGGATTGGAGTCTACCGAGTATGCAAAATGGAGCGT 1320
 Db 1261 aaacatcagcacaagatcttgaaattggattggagctcaccgagtatgcaaatggagcggt 1320
 QY 1321 TTTCTAGCATCTTTCATGAAGGATTAGAGTCAAGTTTCATTTGATTTGATCTGTCCTCT 1380
 Db 1321 ttcttagcatctTTTCATGAAGGATTAGAGTCAAGTTTCATTTGATTTGATCTGTCCTCT 1380
 QY 1381 TTTTGTGTGCTCAAGTTTCTGCTGTTTATGAGTTGCCACTGTTCTTATTTTAAAGTA 1440
 Db 1381 ttttgtgtgctcaagTTTCTGCTGTTTATGAGTTGCCACTGTTCTTATTTTAAAGTA 1440
 QY 1441 TACCTTTGCACTATGTTTGAAGATGCAACGATATAAATGAAATAAATAAATAAATAA 1500
 Db 1441 tacctttgcaCTATGTTTGAAGATGCAACGATATAAATGAAATAAATAAATAAATAA 1500
 QY 1501 AAAAAAA 1508
 Db 1501 aaaaaaa 1508

RESULT 2

AAT37308

ID AAT37308 standard; cDNA to mRNA; 1703 BP.

XX AAT37308;

XX AC

DT 06-FEB-1997 (first entry)

XX Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation; anthocyanin pigment;

XX plants; acylation; colour; tone; colouration; colour change;

XX Gentiana triflora; Petunia hybrida; perilla ocimoides;

XX Senecio cruentus; Lavandula angustifolia; ds.

XX OS Gentiana triflora var. japonica (Clone pGAT4).

XX FH Key Location/Qualifiers

FT CDS 6..1415

FT /*tag= a

XX /product= Aromatic acyl transferase.

XX W09625500-A1.

XX PN 22-AUG-1996.

XX PD 16-FEB-1996; 96WO-JP00348.

XX PF 30-JAN-1996; 96JP-0046534.

XX PR 17-FEB-1995; 95JP-0067159.

XX PR 29-JUN-1995; 95JP-0196915.

XX (SUNR) SUNTORY LTD.

XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;

PI Nakao M, Tanaka Y, Yonekura K;

XX WPI; 1996-393401/39.

DR P-PSDB; AAW04722.

XX

PT DNA coding for aromatic acyl transferase - for transforming plants

PT which produce anthocyanin pigments and thus altering colour tone,

PT e.g. of flowers

XX

PS Claim 4; Page 53-57; 94pp; Japanese.

XX

CC Vectors containing DNA fragments encoding proteins of plant origin

CC with aromatic acyl transferase activity may be used to transform

CC plants which produce anthocyanin pigments. The aromatic acyl

CC transferase acylates the pigments in the flower resulting in colour

CC tone changes and allowing new colourations to be produced. Six

CC specific DNA sequences encoding aromatic acyl transferase from

CC different plants are described in AAT37308-T37313.

XX

XX Sequence 1703 BP; 512 A; 353 C; 356 G; 482 T; 0 other;

XX

Query Match 13.6%; Score 205.6; DB 17; Length 1703;

Best Local Similarity 52.2%; Pred. No. 3.7e-39;

Matches 722; Conservative 0; Mismatches 614; Indels 48; Gaps 10;

QY 1 TGAACATTCTCGAACATGCCGAATATCGGCCCCCTCGGGCACCA---TCGGCCATCGCT 57

DB 25 tgaaggctcttgaaaaatgccaaagttacaccaccatctgacacacagatgctcgagttat 84

QY 58 CGTTATCTCTTACTTCTTCGACATTACTTGGCTACTCTTCCCTCCGGTCCACCATCTTT 117

DB 85 cgtaccggttaacattcttcgatctccctcggttgctgactgaaagaagtcagtcacctc 144

QY 118 TCTTCTATGACTTTCACATTTCTAAATCCCATTTCTATGGACACTATTGTTCCGAGCTAA 177

DB 145 tcttttaacgaacttcgtaaccacgaacacattcttggacactgtataccctaaactcta 204

QY 178 ACAATCTTTATGGTCACTCTTTCACATTTTTCCTCCGTTGCTAGTAATTTGATTTGAT 237

DB 205 agccctcttctctcactcctaaacactacacctccgcttagcggaatttggatgc 264

QY 238 TTCCTAACACTGATGTTTGGGTTTAAATAAANAACCAAAATAAACACAGTTGAAGGTG 297

DB 265 cgaatca-----atcgggcgaaatgccgaagtttcagtaactcccgctgtagggcg 315

QY 298 ATTCTCTGTGTTGTTTCTTTCGAGAATGTTGCTTGACTTTTAATAATTTGACAGGAATC 357

DB 316 actcgataactttgactcgttgcggagctcgcaccagagatttggactaccttaagggtcac 375

QY 358 ATCCTCGAANAATGAAAACCTTTTATCCACTTGTACCTTATTTGGGAAATGCAATCAAT 417

DB 376 aactggtagattcccaatgatttgcatggcctttttatgtatgctgcacgggttataaagga 435

QY 418 ---TATGTGATTGCGTCACSGTCCCACTTTTTCACTTCAAGTGACGTTTTCGCGGCT 474

DB 436 ccatgcaagactataaaagtgtaccgcgtgtagccgtgcaagtaaccgttttccctaac 495

QY 475 CGGTATATCATAGGAATGACGAATCATCATAGCCCTTGGTACGCTAGCACGCGTTCA 534

DB 496 gtggcatagccgtggtctctgacggcacatcattcaatgacagatgctaaaaagtttggtaa 595

QY 535 ACTTTTGAAGGGTGGAGCTTCGATTATTCATTCGATGCTGGTGTAGATCGTCTTTTAAACA 594

DB 556 tgttcataatgcttgggacctatttaacaaatttgggaagacgcgagctgtgtccg 615

QY 595 AAGGATCTCCACGGGTTTGTATAGATTGATTAAACA-----TCCACATTTACATGA 646

DB 616 cgaatctcttccatcttccgatagatcataatcaaatgctgtatggccctagagagaa 675

QY 647 AATAAGTTTGAGACATACAAAG----GCTCGAAAGTTTATTAACCTTCGAGCCTTTGTTG 702

DB 676 cattttggaacgaatgcaagatgttcttgaaatgttctctagatttggaaagcaaacccc 735

QY 703 GTCCCACTGATAAAGTTCGCTCAACGTTTGTGTGACCCGAACTAATATCAATCTACTAA 762

DB 736 ctcgattcaacaaggtaacgagctacatgtctctctcttctgctgaaatccagaagctaa 795

QY 763 AGAAAAGCTCTTAA---CCCAAGTGCCANAATTTGGAGTACATGT--CATCTTTTACGG 816

DB 796 agaacaagtaactgaaatctcagaggatccgaacgacataacgtgtaacgacgttccaa 855

QY 817 TAACCTTGTTGTTATATATGGAGTTGCATAGCGAAATCA-----CTCGTAAAAATAG 867

DB 856 tgacgtgtgatacgtatgacatgcatgctcaaatcaaaagatgacgtcgatcagagg 915

QY 868 GAGAAAGAAAGGGCGAAGAGGAGTTAGAACAGTTTCAATATCACCATTGATGTCGATCTC 927

DB 916 aatcatcgacgacgacgaaatgagctcagctactctcttctacagcgaggtcgcgagac 975

QY 928 GTCTTCATCCCAATTCACAGCCCTACTTTGGTAACCTGTGGTGACCATGTGTCGCCGA 987

DB 976 tctgacgcccccgctgctccgcttaactcttggcaactgtcttgcgtcatgcgttgcga 1035

QY 988 CCTTAAAAAATGCTGTTTGTGACTTACGGAATAATGGGTATGCACCTTGTGCTAAAGTAATTG 1047

DB 1036 aagcaacacataaaagagttagtgtgggataaaaggcttctgttcagttgcagctatg 1095

QY 1048 GAGAGTCTATATGCAAAATGATATATAAAGGACCGGAATCTTTGAAAGATGCCCGGAGAT 1107

DB 1096 gagaagccattgaaaagaggttgacacgaaaaagcgcttcttgcagatgcaaaaact 1155

QY 1108 GG--CATGAACCTTTCATGATCCCGGCTAGGAAGATT---GGTGTGCTGGTACACCTA 1161

DB 1156 ggtatcggaacttaataatgcccttcaaaaagattcttcggtgataccggtatcgcccta 1215

QY 1162 AGCTCAACTTGTACGACTTTGATTTTGGGTGGGGAAGCGCATAAAGATGATGAGACTGTTT 1221

DB 1216 agtctgattcgatggtgtagatttttgatgggaaagcctgcataatttgacattacct 1275

QY 1222 CAATAGACTATATAGCTGATTTCTTATATAATGCAAGCAAAACATCAGCACAGATCTTG 1281

DB 1276 cgtgtgattatgcagaattgatttatgtattcagtcgaggggatttttgaagaaggtg 1335

QY 1282 AATTCGATGAGTCTACCGNGTATGCAATGGAGCGGTTTCTAGCATCTTTGATGAAG 1341

DB 1336 agattggagatacatgctcctaaagattcatggtcatttgcataaaatcttgaagaag 1395

QY 1342 GATT 1345

DB 1396 gctt 1399


```

QY 1132 CTA---GGAAGATGGTGTGCTGTACACCTAAGCTCAACTGTACGACTTTGATTTTG 1188
D 1123 aaactgtttattctcgtggcggtatcgagcagcgttgatctttacgcgcggtatttg 1182
D 1189 GTGGGGGGAAGCGCATAAAGTATGAGCTGTTTCATAGACTATATACGT---CGATTT 1245
D 1183 gatggggtaagcggtgaagcaagagactctgctgattgatgagagaagttaacagtg 1242
QY 1246 CTATAAATGCAAGCAACATCAGCACAGATCTTCAAAATTCGATGTACCGAGTA 1305
D 1243 cgttctgtaaacccgaggtatcgcgaggtatggaggttgattcttcttgccaaagg 1302
QY 1306 TCAAAATGAGCGGTTTCTAGCATCTTTGATGAAGGATTAGAGAGTCAAGTTTCATTTG 1365
D 1303 aggaattgcaagctttgatgatatttttgcgaggggaataaagggtgatttaacattt 1362
QY 1366 AGATCATGTCCTCCCTTTTGTGTGTCATCAAGTTTCTGCGTTTATGAGTGGCACGT 1425
D 1363 a-----atcatgtattatgaagttggatgaatcctctgtttcatctctattgtt 1412
QY 1426 TCTATTCTTTAAGTATACCTTTTCGACTATGTTTTCAGAGATGCAACGATATAAATGA 1485
D 1413 taacaataatttttccattgaaacttttttgagccaataaaaaaaaaaaaaaaaaa 1472
QY 1486 AAAAAAAAAAAAAA 1498
D 1473 aaaaatgaaaaa 1485

RESULT 5
AAT37309
ID AAT37309 standard; cDNA to mRNA; 1622 BP.
AC AAT37309;
XX
DT 06-FEB-1997 (first entry)
DE
DE Aromatic acyl transferase coding sequence.
XX
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KW plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
KW Scenecio cruentus; Lavandula angustifolia; ds.
XX
OS Gentiana triflora var. japonica (Clone pGAT106).
PH
PH Key Location/Qualifiers
FT CDS 6..1415
FT FT /*tag= a
FT FT /product= Aromatic acyl transferase.
FT FT misc_feature 701..703
FT FT /*tag= b
FT FT /transl_except= ATC encodes Leu
FT FT misc_feature 1322..1324
FT FT /*tag= c
FT FT /transl_except= AAT encodes Asp
XX
XX WO9625500-A1.
XX
XX 22-AUG-1996.
XX
XX 16-FEB-1996; 96WO-JP00348.
XX
XX 30-JAN-1996; 96JP-0046534.
XX
XX 17-FEB-1995; 95JP-0067159.
XX
XX 29-JUN-1995; 95JP-0196915.
XX
XX (SUNR ) SUNTORY LTD.
XX
XX Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
XX Nakao M, Tanaka Y, Yonekura K;
XX

```

```

DR WPI; 1996-393401/39.
XX P-PSDB; AAW04723.
PT DNA coding for aromatic acyl transferase - for transforming plants
PT which produce anthocyanin pigments and thus altering colour tone,
XX e.g. of flowers
XX Claim 4; Page 57-61; 94pp; Japanese.
XX Vectors containing DNA fragments encoding proteins of plant origin
CC with aromatic acyl transferase activity may be used to transform
CC plants which produce anthocyanin pigments. The aromatic acyl
CC transferase acylates the pigments in the flower resulting in colour
CC tone changes and allowing new colourations to be produced. Six
CC specific DNA sequences encoding aromatic acyl transferase from
CC different plants are described in AAT37308-T37313.
XX
SQ Sequence 1622 BP; 478 A; 354 C; 313 G; 477 T; 0 other;

Query Match          9.1%; Score 136.6; DB 17; Length 1622;
Best Local Similarity 48.5%; Pred. No. 6e-23;
Matches 675; Conservative 0; Mismatches 644; Indels 72; Gaps 8;

QY 1 TCAACATTCTCGAATCCCGCAATATCGGCCCTCGGCGCACCATCGCCATCGCTCGT 60
D 57 tcaagttcttgagaaatgcggtgttgcgccaccacgcgcgtcgcgagtttacag 116
QY 61 TATCTCTTACTTCTTCGACATTACTTGGCTACTTTCCTCCGGTCCACCATCTTTTCT 120
D 117 tccactgcgttttccgacatcgatggttgcctctgatgcagaacaccatctgcatt 176
QY 121 TCTATGACTTTCCACATTCTAAATCCCATTTTCATGGACACATATTCTCCAGGCTAA 180
D 177 tctacagattccgcacattctgtcccaactctaaatttatcatctccatcca---ttaa 233
QY 181 AATCTTTATCGGTCACTCTTCAACATTTTTTTCCCGTTGCTAGTAATTTGATTTATTC 240
D 234 cgtcccttccctgttctcacaacatttctcgttagccggaatttggattggccgg 293
QY 241 CTACACGTGATGGTTCGGGTTTTTAATAAAACACGAGAATAAACAACGTTGAAGTGAT 300
D 294 tagattctctccg-----atagaatgcggagttgctgttacaagaaggggact 341
QY 301 CTGTTGTGGTTACTTTTGCAGAAATGTTCTGTGACTTTTAATAAATTTGACAGAAATCATC 360
D 342 cgtttctttaaacaattgcagaatccacagatggatttgattatctcgcgagagatcac 401
QY 361 CTCGAAATGTGAAACATTTTATCCACTTTGACCTTCATTGGGAAATGCAATCAAAATAT 420
D 402 agaggattcttataaattcaacgatttgcctccagctccccaagaccattgttaacct 461
QY 421 GTGATTCGCTCAGGTCCTCCACTTTTTCACCTCAAGTCACGCTTTTTCGGGCTCGGTA 480
D 462 ccggcagcagaagtattaccacttttgccttcacaggtgacggtgttctccaacccgta 521
QY 481 TATCATTAGGAATGACGAATCATCATAGCTTGTGTGGCTAGCAGCTAGCAGCGGTTC 540
D 522 tatgcatggagcgaatctctcatcaagttcttggtagccagttcttcttgcatttta 581
QY 541 TGAAGGTTGGACTTCGATTATTCAAATCTGTGTAGAT-----CGGT 582
D 582 ataaattatgggtttgttgacaaatccaatggagattcattaaagtctctccacttt 641
QY 583 CTTTTTTAACGAAGGATCTCCACCGGTTTTTGTAGATTGATTAAACATCCACATTTAG 642
D 642 ctctctacctatgtacgacagatctgtggtgcaagaatccatttccattctcgtcgaa 701
QY 643 ATGAAATTAAGTTGACATACAAAGGCTCGAAAGTTTTTATAAATCTCGAGCTTTGTTG 702
D 702 tctacaatgaaagaaactgctcaaatctcagggcacacctactgttcttaaatccagaa 761
QY 703 GTCCACTGATAAAGTTTCGGTCAACGTTTGTGTGACCCGAACTAATATCAATCTACTAA 762

```


[illegible]


```

PR 26-JUL-1999; 99US-0145695.
XX 17-MAR-2000; 2000US-0190259.
PA (CLIN-) CLINICAL MICRO SENSORS INC.
XX Umek RM;
PI WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 938 BP; 4 A; 144 C; 9 G; 5 T; 776 other;
SQ

Query Match 8.1%; Score 121.6; DB 22; Length 938;
Best Local Similarity 0.9%; Pred. No. 1.7e-19;
Matches 7; Conservative 474; Mismatches 283; Indels 0; Gaps 0;

QY 114 CTTTTCTCTATGACTTTCCACATTTCTAAATCCCATTTTCATGACACTATTTGTTCCAGG 173
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
788 CMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 729

QY 174 CTAACAACATCTTATCGGTCACTCTTCAACATTTTTCGCCGTTTCTGCTAGTAATTCATT 233
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
728 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 669

QY 234 GTATTTCCTAACACTGATGTTCCGGCTTTTAAATAAAACCCAGAAATAAACACGTTGAA 293
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
668 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 609

QY 294 GGTGATCTGTGTTGTTACTTTTGAGAAATGTTGTTGACTTTTAAATATTTGACAGGA 353
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
608 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 549

QY 354 AATCATCCTCGAAATGCGAAACCTTTTATCCACTTTGTACCTTCATGGAATGCAATC 413
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
548 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 489

QY 414 AAATTATGTGATTCGCTCACGGTCCACACTTTTTCACCTCAAGTACGCTTTTTCGGGC 473
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
488 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 429

QY 474 TCGGGTATACATAGGAATGAGGAATCATCATAGCCTTGGTGACGCTAGCACGCGTTC 533
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
428 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 369

QY 534 AACTTTTGAAGGGTGGACTTCGATTATCAATCTGTTGATGATCGGCTTTTAAAG 593
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
368 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 309

QY 594 AAAGGATCCACCGCTTTTGTAGATGATTAAACATCCACATTTAGATGAAATAAG 653
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
308 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 249

QY 654 TTGAGACATACAGGCTCGAAAGCTTTTATAAACCTTCGAGCCTTGTGTCCTCACTGAT 713
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
248 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 189

QY 714 AAAGTTCGGTCAACGTTTGTGTGACCCGAACTAATATCAATCTACTAAGAAAGGTC 773
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
188 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 129

```

```

QY 774 TTAACCAAGTGCCAAACTTGAGTACATGTCTTTTACGTAACATCTGTGTTATATA 833
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
128 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 69

QY 834 TGGAGTTGCATACGGAATCAGTCTGTAATAATAGGAGAAAGAA 877
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
68 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 25

RESULT 12
AAF58252
ID AAF58252 standard; DNA; 936 BP.
XX
XX AAF58252;
XX
XX 24-APR-2001 (first entry)
XX
XX Oligonucleotide D1835.
XX
XX Electron-transfer group. ETM; mismatch; genotyping;
KW gene expression; ss.
XX
XX Synthetic.
XX
XX WO200107665-A2.
XX
XX 01-FEB-2001.
XX
XX 26-JUL-2000; 2000WO-US20476.
XX
XX 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX
XX Umek RM;
XX
XX WPI; 2001-159728/16.
XX
XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX
XX Example 6; Page 127; 159pp; English.
XX
XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.
XX
XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;
SQ.

Query Match 7.8%; Score 118.2; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 1.1e-18;
Matches 6; Conservative 469; Mismatches 282; Indels 0; Gaps 0;

QY 112 ATCTTTCTCTATGACTTTCCACATTTCTAAATCCCATTTTCATGACACTATTTGCCCA 171
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
32 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 91

QY 172 GGCTAAACAATCTTATCGTCACCTCTTCAACATTTTTCGCCGTTGCTAGTAATTTGA 231
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
92 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 151

QY 232 TTCTATTCTCAACATGATGGTTCGGGTTTAAATAAAACCCAGAAATAAACACGTTG 291
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :
152 WMM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM:::MM 211

QY 292 AAGGTGATCTGTGTTGTTTACTTTTGCAGAAATGTTGCTTGAATAATTTGACAG 351
Db |::: : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db 212 271
Qy 352 GAAATCATCTCGAAATGTGAAATCTTTATCCACTTGTACCTTCATTTGGGAAATGCAA 411
Dy 331
Qy 412 TCAAAATATGATGCGTACGAGTCCGACCTTTTTCACCTTCAAGTGACGTTTTCCTGG 471
Dy 391
Qy 472 GCTCGGTATATCATAGGAATGAGAAATCATCATAGCTTGGTGACGTACGACGCGT 531
Dy 451
Qy 532 TCAACTTTTGAAGGGTGGAGCTTCGATTAATCAATCTGGGTAGATCGGCTTTTAA 591
Dy 511
Qy 592 CGAAAGGATCTCCACCGGTTTTTGTAGATTGATTAAACATCCACATTTAGATGAAATA 651
Dy 571
Qy 652 AGTTGAGACATACAGGCTCGAAATGTTTATAACCTTCGAGCCTTGTGTCCTCCACTG 711
Dy 631
Qy 712 ATAAAGTTTCGGTCAACGTTTGTGTTGACCGCACTTAATATCAATCTACTAAAGAAAGG 771
Dy 691
Qy 772 TCTTAACCAAGTGCCAACTTGGAGTACATGTCATCTTTTACGGTAACCTTGTGTTATA 831
Dy 751
Qy 832 TATGAGTTGATACCGGAAATCACTCGTAAATAAGG 868
Dy 788

RESULT 13

AAF58254
ID AAF58254 standard; DNA; 936 BP.

XX AC AAF58254;

XX DT 24-APR-2001 (first entry)

XX DE Oligonucleotide D1875.

XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

XX OS Synthetic.

XX PN WO200107665-A2.

XX PD 01-FEB-2001.

XX PF 26-JUL-2000; 2000WO-US20476.

XX PR 26-JUL-1999; 99US-0145695.

XX PR 17-MAR-2000; 2000US-0190259.

XX PA (CLIN-) CLINICAL MICRO SENSORS INC.

XX PI Umek RM;

XX DR WPI; 2001-159728/16.

XX Nucleic acids containing electron-transfer group, useful as labels in
PT hybridization assays, e.g. for genotyping, allowing repeat analyses on
PT a single surface -
XX

PS Example 6; Page 127; 159pp; English.

XX The present invention relates to a composition comprising two nucleic
CC acids each containing an electron-transfer group (ETM) having
CC different redox potentials. The invention is used for electronic
CC detection of nucleic acids, especially of substitutions (mismatches)
CC and single-nucleotide polymorphisms, e.g. for genotyping,
CC monitoring gene expression.

XX SQ Sequence 936 BP; 4 A; 144 C; 7 G; 5 T; 776 other;

Query Match 7.8%; Score 118.2; DB 22; Length 936;
Best Local Similarity 0.8%; Pred. No. 1.1e-18;
Matches 6; Conservative 469; Mismatches 28; Indels 0; Gaps 0;

Qy 112 ATCTTTCTTCTATGACCTTTCACATTTCTAAATCCATTTTCATGGACACTATTTGCCA 171
Dy 91
Qy 172 GGCTAAACAATCTTTATCGGTCACCTTCAACATTTTTCCTCGTTTGTAGTAATTTGA 231
Dy 151
Qy 232 TTGATTTCTTAACACTGATGGTTCGGGTTTTTAATAAAACACAGAAATAAACACGTTG 291
Dy 211
Qy 292 AAGGTGATCTGTGTGTACTTTTGCAGAAATGTTGCTTGTACTTTAATATTTGACAG 351
Dy 271
Qy 352 GAAATCATCTCGAAATGTGAAACCTTTTATCCACTTTGTACCTTCATTTGGGAAATGCAA 411
Dy 331
Qy 412 TCAAAATATGATGCTGCTACGCTCCACTTTTTCACCTCAAGTGACGTTTTCCTGG 471
Dy 391
Qy 472 GCTCGGTATATCATAGGAATGAGAAATCATATAGCTTGGTGACGCTAGCAGCGGT 531
Dy 451
Qy 532 TCAACTTTTTCGAAAGGTGACTTCGATTATCAATCTGCTAGATCGGCTTTTAA 591
Dy 511
Qy 592 CGAAAGGATCTCCACCGGTTTTTGTAGATTGATTAAACATCCACATTTAGATGAAATA 651
Dy 571
Qy 652 AGTTGAGACATACAGGCTCGAAAGTTTTTATAACCTTCGAGCCTTGTGTCCTCCACTG 711
Dy 631
Qy 712 ATAAAGTTTCGTCACGCTTTCGTTGACCGCACTTAATATCAATCTACTAAAGAAAGG 771
Dy 691
Qy 772 TCTTAACCAAGTGCCAACTTGGAGTACATGTCATCTTTTACGGTAACCTTGTGTTATA 831
Dy 751
Qy 832 TATGAGTTGATACCGGAAATCACTCGTAAATAAGG 868
Dy 788

RESULT 14

AAF58257
ID AAF58257 standard; DNA; 936 BP.

XX

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:04:45 ; Search time 168.74 Seconds
(without alignments)
1691.841 Million cell updates/sec

Title: US-08-894-356C-5
Perfect score: 1508
Sequence: 1 TGAACATTCCTCGAACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA: *
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq: *
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq: *
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq: *
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq: *
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq: *
6: /cgn2_6/ptodata/2/ina/backfiles1.seq: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	53.6	3.6	7218	1	US-08-232-463-14 Sequence 14, Appl
2	47	3.1	1472	4	US-08-781-420-10 Sequence 10, Appl
3	47	3.1	1472	4	US-08-781-420-12 Sequence 12, Appl
4	46.6	3.1	1534	1	US-08-300-903A-6 Sequence 6, Appl
5	44.6	3.0	1641	1	US-08-300-903A-8 Sequence 8, Appl
6	43.4	2.9	1493	6	5340934-5 Patent No. 5340934
7	42.8	2.8	1098	3	US-09-248-335-35 Sequence 35, Appl
8	42.6	2.8	2589	4	US-08-569-749-1 Sequence 1, Appl
9	42.6	2.8	2589	5	PCT-US96-12860-1 Sequence 1, Appl
10	42.2	2.8	1223	3	US-09-154-874-4 Sequence 4, Appl
11	42	2.8	1474	4	US-08-821-994-64 Sequence 64, Appl
12	41.8	2.8	991	3	US-08-924-747-25 Sequence 25, Appl
13	41.8	2.8	991	3	US-09-247-373B-25 Sequence 25, Appl
14	41.8	2.8	991	4	US-09-296-715-25 Sequence 25, Appl
15	41.8	2.8	1700	2	US-08-897-340-4 Sequence 4, Appl
16	41.8	2.8	1700	3	US-09-252-323-4 Sequence 4, Appl
17	41.6	2.8	893	1	US-08-276-452A-66 Sequence 66, Appl
18	41.6	2.8	893	2	US-08-798-744-66 Sequence 66, Appl
19	41.6	2.8	2030	3	US-08-706-216-3 Sequence 3, Appl
20	41.6	2.8	3319	3	US-08-335-844A-15 Sequence 15, Appl
21	41.4	2.7	2744	3	US-09-071-101-1 Sequence 1, Appl
22	41.4	2.7	2744	3	US-09-369-618-1 Sequence 1, Appl
23	41.4	2.7	2744	3	US-09-369-617-1 Sequence 1, Appl
24	41	2.7	1325	1	US-08-306-691B-51 Sequence 51, Appl
25	41	2.7	1325	2	US-08-464-517-1 Sequence 1, Appl
26	41	2.7	1325	2	US-08-246-361A-1 Sequence 1, Appl
27	41	2.7	1325	3	US-08-463-772-1 Sequence 1, Appl

28 41 2.7 1325 5 PCT-US93-05000-1 Sequence 1, Appl
29 41 2.7 1577 4 US-08-821-994-59 Sequence 59, Appl
30 40.8 2.7 350 1 US-08-171-385-14 Sequence 14, Appl
31 40.8 2.7 350 3 US-08-361-441B-14 Sequence 14, Appl
32 40.8 2.7 1728 3 US-08-985-950-7 Sequence 7, Appl
33 40.8 2.7 1738 2 US-08-379-482A-2 Sequence 2, Appl
34 40.8 2.7 8920 2 US-08-446-855A-1 Sequence 1, Appl
35 40.8 2.7 8920 4 US-09-150-741-1 Sequence 1, Appl
36 40.6 2.7 1166 5 PCT-US96-12129B-1 Sequence 10, Appl
37 40.6 2.7 1582 3 US-08-545-196B-12 Sequence 12, Appl
38 40.6 2.7 1582 3 US-08-545-196B-12 Sequence 12, Appl
39 40.6 2.7 5173 1 US-08-242-677-1 Sequence 1, Appl
40 40.4 2.7 1542 1 US-08-207-904-1 Sequence 1, Appl
41 40.4 2.7 3706 1 US-08-207-904-16 Sequence 16, Appl
42 40.2 2.7 340 1 US-08-171-385-27 Sequence 27, Appl
43 40.2 2.7 340 3 US-08-361-441B-27 Sequence 27, Appl
44 40 2.7 1075 4 US-08-400-006B-6 Sequence 6, Appl
45 40 2.7 1332 2 US-09-057-762-1 Sequence 1, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pT29pt-fls
US-08-232-463-14

Query Match 3.6%; Score 53.6; DB 1; Length 7218;
Best Local Similarity 1.9%; Pred. No. 0.00017;
Matches 5; Conservative 171; Mismatches 90; Indels 0; Gaps 0;

Qy 7 TTCGCAACATGCCGATATCGGCCCTCGGGCACCATCGGCATCGTCGTTATCTC 66
Db 1177 YY 1236
Qy 67 TTACTTTCTCGACATTACTTGGCTACTCTCTCCCTCCGGTCCACCATCTTTTCTCTATG 126
Db 1237 YY 1296
Qy 127 ACTTTCCACATCTAAATCCCATTTCAFGACATATTTGCCAGGTAAACAATCTT 186
Db 1297 YY 1356
Qy 187 TATCGGTCACTCTCAACATTTTTCCTCGTTGTAGTAATTTGATTTCTTCTAACA 246
Db 1357 YY 1416
Qy 247 CTGATGGTTCGGTGTAAATAAAAAA 272
Db 1417 YYYYYYYYYYYYYYYYYYGTACCAA 1442

RESULT 2
US-08-781-420-10
; Sequence 10, Application US/08781420
; Patent No. 6248872
; GENERAL INFORMATION:
; APPLICANT: Chandrashekar, Ramaswamy
; APPLICANT: Mehta, Kapil
; TITLE OF INVENTION: Parasitic Nematode Tranglutaminase
; TITLE OF INVENTION: Proteins, Nucleic Acid Molecules and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08781.420
; FILING DATE: December 3, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1472 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2..1105
US-08-781-420-10

Query Match 3.1%; Score 47; DB 4; Length 1472;
Best Local Similarity 56.0%; Pred. No. 0.0048;
Matches 89; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 1344 TTAGAGAGTCAAGTTTCATTAGATCATCGTCCCTTTTGTGTGATCAAGTTTCTGT 1403
Db 1314 TCAGTTCGGAATGTTTATTCGGTTAGCTTAGGCTTTTTTTTGTATGTTACTGT 1373
Qy 1404 CGTTTTATGAGTGGCCACTGTTCTATTCTTTTAAGTATACCTTTCGACTATGTTTGAAG 1463
Db 1374 TGTATTGTATTACTATTTTCCCTTGTGTTTTAAATTTAAATAAAATTTTTTTTGGAA 1433
Qy 1464 ATGCAACGATATAAATGAAAAAATAAAAAA 1502
Db 1434 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1472

RESULT 3
US-08-781-420-12/c
; Sequence 12, Application US/08781420
; Patent No. 6248872
; GENERAL INFORMATION:
; APPLICANT: Chandrashekar, Ramaswamy
; APPLICANT: Mehta, Kapil
; TITLE OF INVENTION: Parasitic Nematode Tranglutaminase
; TITLE OF INVENTION: Proteins, Nucleic Acid Molecules and
; TITLE OF INVENTION: Uses Thereof
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Carol Talkington Verser, Ph.D.
; ADDRESSEE: Heska Corporation
; STREET: 1825 Sharp Point Drive
; CITY: Fort Collins
; STATE: Colorado
; COUNTRY: USA
; ZIP: 80525
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: Windows 95
; SOFTWARE: WordPerfect for Windows, Version 7.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08781.420
; FILING DATE: December 3, 1996
; CLASSIFICATION: 536
; ATTORNEY/AGENT INFORMATION:
; NAME: Verser, Carol Talkington
; REGISTRATION NUMBER: 37,459
; REFERENCE/DOCKET NUMBER: HW-2
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 970/493-7272
; TELEFAX: 970/484-9505
; INFORMATION FOR SEQ ID NO: 12:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1472 nucleotides
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
US-08-781-420-12

Query Match 3.1%; Score 47; DB 4; Length 1472;
Best Local Similarity 56.0%; Pred. No. 0.0048;
Matches 89; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

Qy 1344 TTAGAGAGTCAAGTTTCATTAGATCATCGTCCCTTTTGTGTGATCAAGTTTCTGT 1403
Db 159 TCAGTTCGGAATGTTTATTCGGTTAGCTTAGGCTTTTTTTTGTATGTTACTGT 100
Qy 1404 CGTTTTATGAGTGGCCACTGTTCTATTCTTTTAAGTATACCTTTCGACTATGTTTGAAG 1463
Db 99 TGTATTGTATTACTATTTTCCCTTGTGTTTTAAATTTAAATAAAATTTTTTTTGGAA 40

QY 1464 ATCCACGATATATAATGAAAAAAGAAAAAAGAAAAAAGAAAAA 1502
 Db 39 AA 1

RESULT 4
 US-08-300-903A-6
 ; Sequence 6, Application US/08300903A
 ; Patent No. 5591630
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, Dirk M
 ; APPLICANT: Giri, Judith G
 ; TITLE OF INVENTION: Interleukin-15 Receptors
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Immunex Corporation
 ; STREET: 51 University Street
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Apple Operating System 7.1
 ; SOFTWARE: Microsoft Word for Apple, Version 5.1a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/300,903A
 ; FILING DATE: 06-SEPTEMBER-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: USSN 08/236,919
 ; FILING DATE: 06-MAY-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Perkins, Patricia Anne
 ; REGISTRATION NUMBER: 34,695
 ; REFERENCE/DOCKET NUMBER: 2822-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-587-0430
 ; TELEFAX: 206-233-0644
 ; INFORMATION FOR SEQ ID NO: 6:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1534 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..753
 ; US-08-300-903A-6

Query Match 3.1%; Score 46.6; DB 1; Length 1534;
 Best Local Similarity 60.8%; Pred. No. 0.0061;
 Matches 93; Conservative 0; Mismatches 59; Indels 1; Gaps 1;
 QY 1357 TTTTCATTGATCATCGTCCCTTTTGTGTCATCAAGTTTCTGTCGTTTATGAGT 1416
 Db 1368 TATATTTATCTCTGTCCTATTTATCTATATCTATATATATTTGTTGTCAT 1427
 QY 1417 TGGCAGTGTCTATCTTTTA-AGTATACCTTTTCGACTATGTTTGAAGATGCAACGAT 1475
 Db 1428 TGTACCTTGTATACAAAAATAAAACATCTATTTTCAATATTTTAAATGCAAAAAA 1487
 QY 1476 AAAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1508
 Db 1488 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1520

RESULT 5
 US-08-300-903A-8
 ; Sequence 8, Application US/08300903A
 ; Patent No. 5591630
 ; GENERAL INFORMATION:
 ; APPLICANT: Anderson, Dirk M
 ; APPLICANT: Giri, Judith G
 ; TITLE OF INVENTION: Interleukin-15 Receptors
 ; NUMBER OF SEQUENCES: 15
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Immunex Corporation
 ; STREET: 51 University Street
 ; CITY: Seattle
 ; STATE: Washington
 ; COUNTRY: USA
 ; ZIP: 98101
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: Apple Macintosh
 ; OPERATING SYSTEM: Apple Operating System 7.1
 ; SOFTWARE: Microsoft Word for Apple, Version 5.1a
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/300,903A
 ; FILING DATE: 06-SEPTEMBER-1994
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: USSN 08/236,919
 ; FILING DATE: 06-MAY-1994
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Perkins, Patricia Anne
 ; REGISTRATION NUMBER: 34,695
 ; REFERENCE/DOCKET NUMBER: 2822-A
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 206-587-0430
 ; TELEFAX: 206-233-0644
 ; INFORMATION FOR SEQ ID NO: 8:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1641 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: cDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 3..839
 ; US-08-300-903A-8

Query Match 3.0%; Score 44.6; DB 1; Length 1641;
 Best Local Similarity 56.5%; Pred. No. 0.02;
 Matches 83; Conservative 0; Mismatches 64; Indels 0; Gaps 0;
 QY 1362 TTCTAGATCATCGTCCCTTTTGTGTCATCAAGTTTCTGTCGTTTATGAGTCCCA 1421
 Db 1453 TTATATTATATCTCTGTCCTATTTATATCTATATATATATATATATATTTGTTGACA 1512
 QY 1422 CTGTTCTATTCTTTTAAGTATACCTTTTCGACTATGTTTGAAGATGCAACGATATAAATG 1481
 Db 1513 TTGTACCTTGTATACAAAAATAAAACATCTATTTTCAATATAAAAAAAAAAAAAA 1572
 QY 1482 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1508
 Db 1573 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1599
 RESULT 6
 5340934-5
 ; Patent No. 5340934
 ; APPLICANT: TERMINE, JOHN D.; YOUNG, MARIAN F.; FISHER, LARRY W.


```

? TELEFAX: 302-773-0164
? INFORMATION FOR SEQ ID NO: 4:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 1223 base pairs
? TYPE: nucleic acid
? STRANDEDNESS: single
? TOPOLOGY: linear
? MOLECULE TYPE: CDNA
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 55..1223
? FEATURE:
? NAME/KEY: CDS
? LOCATION: 55..990
? US-09-154-874-4
?
? Query Match 2.8%; Score 42.2; DB 3; Length 1223;
? Best Local Similarity 63.1%; Pred. No. 0.074;
? Matches 65; Conservative 0; Mismatches 38; Indels 0; Gaps 0;

```

```

Db 1107 TTTTCGGATCTTGAAATGCTGAACGAGTTTAACTTTATCATTAATTTAAGCTCTGTTGTT 1166
Qy 1466 GCAACGATATAAAATGAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1508
Db 1167 TTTAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 1209

RESULT 11
US-08-821-994-64
; Sequence 64, Application US,08821994A
; Patent No. 6228643
; GENERAL INFORMATION:
; APPLICANT: Greenland, Andrew J
; APPLICANT: Thomas, Didier EP
; APPLICANT: Jepson, Ian

```

```

: FILE REFERENCE: PED J0100
: CURRENT APPLICATION NUMBER: US/08/821,994A
: CURRENT FILING DATE: 1997-03-22
: EARLIER APPLICATION NUMBER: PCT/GB97/00729
: EARLIER FILING DATE: 1997-03-18
: EARLIER APPLICATION NUMBER: GB 9606062.9
: EARLIER FILING DATE: 1996-03-22
: NUMBER OF SEQ ID NOS: 89
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 64
: LENGTH: 1474
: TYPE: DNA
: ORGANISM: Brassica napus
: US-08-821-994-64

Query Match      2.8%; Score 42; DB 4; Length 1474;
Best Local Similarity 57.7%; Pred. No. 0.088;
Matches 75; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

Qy 1379 CTTTTGTGTCATCAAGTCTCTCGTGTTTTATGAGTTGCCACGTGTTCTATTCTTTAAG 1438
Db 1302 ctttccttcacaaagaagatgattggctgtcttatgattaaagaagatcaatacaaat 1361

Qy 1439 TATACCTTTTCGACATGATGTTTTGAAGATCCACGATATAAAATGAAAAAATAAAAAA 1498
Db 1362 gatattctcttaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1421

Qy 1499 AAAAAAATAA 1508
Db 1422 aaaaaaaaaa 1431

RESULT 12

```

```
US-08-924-747-25
; Sequence 25, Application us/08924747
; Patent No. 6063570
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/924,747
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-08-924-747-25

Query Match 2.8%; Score 41.8; DB 3; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Qy 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Db 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Qy 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Qy 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977

RESULT 13
US-09-247-373B-25
; Sequence 25, Application us/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/296,715
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-09-247-373B-25

Query Match 2.8%; Score 41.8; DB 4; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Qy 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Db 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Qy 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Qy 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977

RESULT 14
US-09-296-715-25
; Sequence 25, Application us/09296715
; Patent No. 6171839
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/296,715
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-09-296-715-25

Query Match 2.8%; Score 41.8; DB 4; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Qy 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Db 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Qy 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Qy 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977
```

```
; PRIOR FILING DATE: 1997-09-05
; NUMBER OF SEQ ID NOS: 56
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 25
; LENGTH: 991
; TYPE: DNA
; ORGANISM: SOYBEAN
US-09-247-373B-25

Query Match 2.8%; Score 41.8; DB 4; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Qy 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Db 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Qy 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Qy 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977

RESULT 14
US-09-296-715-25
; Sequence 25, Application us/09296715
; Patent No. 6171839
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN
; APPLICANT: O'KEEFE, DANIEL
; TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: E.I. DU PONT DE NEMOURS AND COMPANY
; STREET: 1007 MARKET STREET
; CITY: WILMINGTON
; STATE: DELAWARE
; COUNTRY: UNITED STATES OF AMERICA
; ZIP: 19898
; COMPUTER READABLE FORM:
; MEDIUM TYPE: DISKETTE, 3.50 INCH
; COMPUTER: IBM PC COMPATIBLE
; OPERATING SYSTEM: MICROSOFT WORD FOR WINDOWS 95
; SOFTWARE: MICROSOFT WORD VERSION 7.0A
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/296,715
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: FLOYD, LINDA AXAMETHY
; REGISTRATION NUMBER: 33,692
; REFERENCE/DOCKET NUMBER: CL-1108
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 302-892-8112
; TELEFAX: 302-773-0164
; INFORMATION FOR SEQ ID NO: 25:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 991 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; TISSUE TYPE: SOYBEAN
; IMMEDIATE SOURCE:
; CLONE: SSM.PK0067.G5
US-09-296-715-25

Query Match 2.8%; Score 41.8; DB 4; Length 991;
Best Local Similarity 59.8%; Pred. No. 0.087;
Matches 70; Conservative 0; Mismatches 47; Indels 0; Gaps 0;

Qy 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Db 1392 TCAAGTTTCTGCGGTTTATGAGTTGCGACGTGTTCTATTCTTTAAGTATACCTTTTCGAC 1451
Qy 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Db 861 TTAATTGACTGTGATTGGTTGGTATGTTGCTATTTTAACTAAATAAAAGTGTTC 920
Qy 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Db 1452 TATGTTTGAAGATCAACGATATATAAATGAAAAAATAAAAAAAAAAAAAA 1508
Qy 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977
Db 921 AGTTTAAAAAATAAAAAAAAAAAAAAATAAAAAAAAAAAAAA 977
```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:00:56 ; Search time 5816.79 Seconds
(without alignments)
2450.648 Million cell updates/sec

Title: US-08-894-356C-5
Perfect score: 1508
Sequence: 1 TGAACATTCTCGAACATGCC.....AAAAAAAAAAAAAAAAAAAA 1508

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*

1:	gb_est1:*
2:	gb_est2:*
3:	gb_est3:*
4:	gb_est4:*
5:	gb_est5:*
6:	gb_est6:*
7:	gb_est7:*
8:	gb_est8:*
9:	gb_est9:*
10:	gb_est10:*
11:	gb_est11:*
12:	gb_est12:*
13:	gb_est13:*
14:	gb_est14:*
15:	gb_est15:*
16:	gb_est16:*
17:	gb_est17:*
18:	gb_est18:*
19:	gb_est19:*
20:	gb_est20:*
21:	gb_est21:*
22:	gb_est22:*
23:	gb_est23:*
24:	gb_est24:*
25:	gb_est33:*
26:	gb_est34:*
27:	gb_est35:*
28:	gb_est36:*
29:	gb_est37:*
30:	gb_est38:*
31:	gb_est39:*
32:	gb_est40:*
33:	em_estba:*
34:	em_estfun:*
35:	em_esthum1:*
36:	em_esthum2:*
37:	em_esthum3:*
38:	em_esthum4:*
39:	em_esthum5:*
40:	em_esthum6:*
41:	em_esthum7:*
42:	em_esthum8:*
43:	em_esthum9:*

44:	em_esthum10:*
45:	em_esthum11:*
46:	em_esthum12:*
47:	em_esthum13:*
48:	em_esthum14:*
49:	em_esthum15:*
50:	em_esthum16:*
51:	em_esthum17:*
52:	em_esthum18:*
53:	em_esthum19:*
54:	em_esthum20:*
55:	em_esthum21:*
56:	em_esthum22:*
57:	em_esthum23:*
58:	em_esthum24:*
59:	em_esthum25:*
60:	em_esthum26:*
61:	em_esthum27:*
62:	em_esthum28:*
63:	em_estin1:*
64:	em_estin2:*
65:	em_estin3:*
66:	em_estin4:*
67:	em_estin5:*
68:	em_estom1:*
69:	em_estom2:*
70:	em_estov1:*
71:	em_estov2:*
72:	em_estpl1:*
73:	em_estpl2:*
74:	em_estpl3:*
75:	em_estpl4:*
76:	em_estpl5:*
77:	em_estpl6:*
78:	em_estpl7:*
79:	em_estpl8:*
80:	em_estpl9:*
81:	em_estpl10:*
82:	em_estrol:*
83:	em_estro2:*
84:	em_estro1:*
85:	em_estro4:*
86:	em_estro5:*
87:	em_estro6:*
88:	em_estro7:*
89:	em_estro8:*
90:	em_estro9:*
91:	em_estro10:*
92:	em_estro11:*
93:	em_estro12:*
94:	em_estro13:*
95:	em_estro14:*
96:	em_estro15:*
97:	em_estro16:*
98:	em_estro17:*
99:	em_estro18:*
100:	em_estro19:*
101:	em_estro20:*
102:	gb_est25:*
103:	gb_est26:*
104:	gb_est27:*
105:	gb_est28:*
106:	gb_est29:*
107:	gb_est30:*
108:	gb_est31:*
109:	gb_est32:*
110:	gb_est41:*
111:	gb_est42:*
112:	gb_est43:*
113:	gb_est44:*
114:	gb_est45:*
115:	gb_est46:*
116:	gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	130.2	8.6	690	155	BG599447	BG599447 EST504342
2	124.4	8.2	763	155	BG591758	BG591758 EST499600
3	116.4	7.7	570	142	BE922784	BE922784 EST426553
4	110.2	7.3	788	164	BE187619	BE187619 EST336180
5	107.6	7.1	575	142	BE920468	BE920468 EST424237
6	107.2	7.1	487	118	AW616206	AW616206 EST307245
7	107	7.1	606	113	AW221049	AW221049 EST297518
8	105.8	7.0	508	173	BG097054	BG097054 EST461573
9	103.4	6.9	574	142	BE922746	BE922746 EST426515
10	103.2	6.8	583	113	AW221050	AW221050 EST297519
11	99	6.6	591	167	BE434257	BE434257 EST405335
12	98	6.5	564	142	BE921494	BE921494 EST425179
13	92.6	6.1	482	144	BF096979	BF096979 EST400516
14	91.4	6.1	630	155	BG581249	BG581249 EST482982
15	89.2	5.9	497	122	AW930934	AW930934 EST356777
16	87.6	5.8	705	114	AW239055	AW239055 EST305729
17	87.2	5.8	659	155	BG598154	BG598154 EST496832
18	85.8	5.7	403	143	BF053202	BF053202 EST438432
19	85	5.6	542	167	BE436185	BE436185 EST407263
20	84.6	5.6	518	119	AW650280	AW650280 EST328734
21	82.6	5.5	687	111	AW102336	AW102336 sdb6d06.y
22	82	5.4	783	155	BG584708	BG584708 EST486469
23	81.6	5.4	526	166	BE323055	BE323055 NF001A07P
24	80.2	5.3	708	138	BE660849	BE660849 7-C8 Gmax
25	79.4	5.3	676	153	BG455847	BG455847 NF071A12P
26	79.4	5.3	681	151	BF639071	BF639071 NF071A12P
27	79.2	5.3	648	154	BG457806	BG457806 NF034C06P
28	78.8	5.2	408	151	BF643876	BF643876 NF089D05E
29	78.8	5.2	642	153	BG448305	BG448305 NF070A05E
30	76.6	5.1	603	118	AW616119	AW616119 EST296884
31	73.8	4.9	654	151	BF642724	BF642724 NF070B03I
32	73.8	4.9	656	119	AW692527	AW692527 NF056D05S
33	73.8	4.9	677	118	AW586972	AW586972 EST318595
34	72.2	4.8	597	120	AW774948	AW774948 EST334099
35	71.6	4.7	605	30	AV440958	AV440958 AV440958
36	71.6	4.7	649	151	BF643475	BF643475 NF004H05E
37	71.6	4.7	694	151	AW560686	AW560686 EST315734
38	71.6	4.7	829	155	BG584678	BG584678 EST486439
39	71.2	4.7	411	142	BE922378	BE922378 EST426147
40	71.2	4.7	414	142	BE921952	BE921952 EST425721
41	67.6	4.5	433	120	AW756036	AW756036 s112h06.y
42	63.6	4.2	368	142	BE921953	BE921953 EST425722
43	62.8	4.2	307	142	BE920010	BE920010 EST423779
44	62	4.1	360	150	BF520206	BF520206 EST457675
45	62	4.1	456	142	BE923572	BE923572 EST427341

ALIGNMENTS

RESULT 1	BG599447	690 bp	mRNA	EST	12-APR-2001
LOCUS	EST504342	cSTS Solanum tuberosum	cDNA clone	cSTS25122	5' sequence,
DEFINITION	mRNA sequence.				
ACCESSION	BG599447				
VERSION	BG599447.1	GI:13616583			
KEYWORDS	EST.				
SOURCE	potato.				
ORGANISM	Solanum tuberosum				
REFERENCE	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
AUTHORS	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
TITLE	Asteridae; euasterids I; Solanales; Solanaceae; Solanum.				
	1 (bases 1 to 690)				
	van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,				
	Bouiri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.				
	Generations of ESTs from sprouting potato eyes				

JOURNAL COMMENT

Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.

FEATURES

source
1..690
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS25122"
/clone_lib="cSTS"
/issue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"

note=Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."

BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN

Query Match 8.5%; Score 130.2; DB 155; Length 690;
Best Local Similarity 52.3%; Pred. No. 6e-21;
Matches 316; Conservative 0; Mismatches 268; Indels 15; Gaps 1;

QY 21 CGAATATCGGCCCTCGGSCACCATCGGCATCGCTCGTATCTTACTTTCTTCGAC 80
DB 9 CAAGTGGCCGACCTCGGSCGCGGAGGAGGTGACACTCCCTCTTACTTTTGTAT 68
QY 81 ATTTACTTGGCTACTCTTCCCTCGGTCACCATCTTTTCTTCTATGACTTTCACACTT 140
DB 69 CATGTTGGTGGTTCACACCGCTATAGGCGGATATATTCTACAAGCTCCCAATTC 128
QY 141 AATCCCATTTATGACACACTATTGTTCCAGGCTAAACAATCTTTATCGGTCACCTTT 200
DB 129 AAACCGGATTCGTTTCAAAACATATTATCCCTCTTAAAAAATTCACCTCCCTC 188
QY 201 CAACATTTTTCGCTTGTAGTAAATTTGATGTTATTTCTTAACACTGATGCTGGGT 260
DB 189 AAACACTATACGCTTGTAGTGAACGTTGCTCGTCCACTAGATACAAACGG- 241
QY 261 TTTTAATAAAACCCAGAAATAAACACGTTGAAGGTTGCTGTTGGTGTACTTTTGA 320
DB 242 -----ATATCCTGAGTTACATTATGTGACAGGAGATTCTGTATCTGTACTTTTCT 293
QY 321 GAATGTTGCTTGCATTTAATAATTTGACAGAAATCATCTCGAAATGTAACATTTT 380
DB 294 GAGACTGATGATTTCAATTTATCTCATTTGGTGACCATCCGCGGTATGCTAGGATTT 353
QY 381 TATCCACTTGTACTTCAATTTGGGAAATGCAATCAAAATTTATGTTGCTGCGTCCCA 440
DB 354 TATCACTTGTCTTCTAGTTAGGGAACCTAAGATGCACCCGGGTCCTCAATTAGCTCC 413
QY 441 CTTTTTTCACCTTCAAGTGACCTTTTTCGCGGCTCGGGTATATCATAGGAATAGCAAT 500
DB 414 GTCTTAGCCATCAAGTGACACTTTTTCGGAATCTTTGGTGTATCCATTTGTTTCACTAAC 473
QY 501 CATCATGCTTGGTGACGCTAGCACCGGTTCAACTTTTTTGAAGGTTGACTTCGATT 560
DB 474 CATCATGTTGTTGGTATGAGGATAGTAGGGTTTCAATAGGGCGTGGGCTTACT 533
QY 561 ATTCAATCTGCTAGATCGTCTTTTAAACGAAAGATCTCCACCGGTTTGTATAG 619
DB 534 CATAAATTCGGTGGAGATGACAAATTTCTTATCGAATGAGCTAATTCATTTTATGATAG 592

RESULT 2
BG591758

```
LOCUS       BG591758             763 bp    mRNA                    EST             12-APR-2001
DEFINITION   EST499600 P. infestans-challenged leaf Solanum tuberosum cDNA clone
              BPL10A14 5' sequence, mRNA sequence.
ACCESSION    BG591758
VERSION      BG591758.1   GI:13609898
KEYWORDS     EST.
SOURCE       potato.
ORGANISM     Solanum tuberosum
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE    1 (bases 1 to 763)
AUTHORS      Zhang,P., Hernandez,M., Tornqvist,C.-E., Wirtz,U., Loukolanov,A.,
              Rangel,P., Haberland,G.T., Cho,J., Chieningo,A., Bougri,O., Buell
              ,C.R., Ronning,C.M., Helgeson,J. and Baker,B.
              Generation of ESTs from Potato Leaves Challenged with Phytophthora
              infestans, Incompatible Reaction
              Unpublished (2000)
JOURNAL
COMMENT      The Institute for Genomic Research
              For clone info: please contact Research Genetics, Libraries
              Division tel 1-800-711-6195, email cdna@resgen.com
              Seq primer: M13P-R.
FEATURES     Location/Qualifiers
              1..763
                 /organism="Solanum tuberosum"
                 /cultivar="Kennebec"
                 /db_xref="taxon:4113"
                 /clone="BPL10A14"
                 /clone_lib="P. infestans-challenged leaf"
                 /tissue_type="leaf"
                 /dev_stage="6 week old"
                 /lab_host="SOLR"
                 /note="vector: pbluescript SK(-); Site_1: EcoRI; Site_2:
              xhoI; Whole plants were challenged with 450,000
              sporangia/ml P. infestans US-1(US 940501) in Biotron
              (Madison, Wisconsin). Leaf tissue was collected at 1, 2,
              5, 12, and 24 hours post-challenge and frozen in liquid
              nitrogen immediately upon removal. Kennebec plants showed
              no signs of HR. Katakhdin plants (susceptible to P.
              infestans US-1) were used as controls and showed
              infection. NOTE: We cannot exclude the possibility that
              this sequence is actually derived from Phytophthora rather
              than potato."
BASE COUNT   212 a 162 c 156 g 233 t
ORIGIN
Query Match      8.2%; Score 124.4; DB 155; Length 763;
Best Local Similarity 53.2%; Pred. No. 1.5e-19;
Matches 297; Conservative 0; Mismatches 246; Indels 15; Gaps 1;
Qy 64 CTCCTACTCTTCGACATTACTTGGCTACTCTTCCCTCCGGTCCACCATCTTTCTCTCT 123
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 48 CTCCTACTTATTTTGATCATGTTTGGTATAGGTTCCACCGTATACGGCGGATATATCT 107
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 124 ATGACTTTCCACATCTATAATCCCATTCATGGACACATATTGTTCCCGAGGCTAAACAAAT 183
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 108 ACAAGTCCCATTTCCAAACCCGATTCGTTCAAAACATTTATCTCTACTCTTAAAAATT 167
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 184 CTTTATCGGTCACTCTTCAACATTTTTTCCGGTTTGGCTAGTAATTTGATTTTCTCTA 243
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 168 CACTCTCCCTCACTCTCAACACATACACGCCCTTAGCAGGAAATGTTGCTCGTCACTAG 227
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 244 ACACGTAGGTTCCGGTTTTTAATAAAACCCAGAAATAAAACAGTTGAAGGTATCTG 303
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 228 ATACAAATGG-----ATATCCTGAGTTACGTTATGTGACAGGAATTTCTG 272
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 304 TTGTGGTTACTTTTGCAGAAATGTTGCTCTTGACTTTTAATAATTTGACAGGAATCATCTC 363
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 273 TATCTGTTACTTTTTTGAGACTGATATGATTTCAATTATCTATCTATGTTGACCATCCG 332
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 364 GAAATGTGAAAACTTTTATCCACTTGTACCTTCAATTGGGAAATGCAATCAATATATGTG 423
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

```
Db 333 GTAATGCTAAGGATTTTATCATTCTTGTCTTAAGTTAGGGAACCTTAAGATGCACCCG 392
    || | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 424 ATTGGTCAACGGTCCCACTTTTTCACATTCAAGTGACCTTTTCCGGGCTCGGTATAT 483
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 393 GGGTCCAATTAGCTCCGGTCTTAGCCATTCAAGTGACACTTTTTCCGAATCTTGGTGAT 452
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 484 CACTAGGAATGACGAATCATAGCTTGGTGACGGCTAGCACGGCTTCAACTTTTGA 543
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 453 CCAITGGTTTCACATAACCATCATGTTGGTGGTAGGAGCTACCATAGAGGTTCAATTA 512
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 544 AAGGTGACCTTCGATTAATCAATCTGGTGATGATCGGTCCTTTTAAACGAAGATCTC 603
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 513 AGGCGTGGGCTCTACTCCATAAATTCGGTGGAGATGAACAATTTCTATCGAATGAGTAA 572
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 604 CACCGGTTTTTGGATAGAT 621
    || ||| ||| ||||| |
Db 573 TTCCATTTTATGATAGGT 590
    || ||| ||| ||||| |

RESULT      3
BE922784
LOCUS       BE922784             570 bp    mRNA                    EST             02-OCT-2000
DEFINITION   EST426553 potato leaves and petioles Solanum tuberosum cDNA clone
              CSTB22E2 5' sequence, mRNA sequence.
ACCESSION    BE922784
VERSION      BE922784.1   GI:10448860
KEYWORDS     EST.
SOURCE       potato.
ORGANISM     Solanum tuberosum
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE    1 (bases 1 to 570)
AUTHORS      van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
              Uterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
              ,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
              Generation of ESTs from potato leaves and petioles
              Unpublished (2000)
JOURNAL
COMMENT      Contact: Cathy Ronning
              The Institute for Genomic Research
              For clone request: please contact Research Genetics, Libraries
              Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES     Location/Qualifiers
              1..570
                 /organism="Solanum tuberosum"
                 /cultivar="Kennebec"
                 /db_xref="taxon:4113"
                 /clone="cSTB22E2"
                 /clone_lib="potato leaves and petioles"
                 /tissue_type="leaflets and petioles"
                 /dev_stage="8 weeks old plants"
                 /lab_host="SOLR"
                 /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
              xhoI; Tissue was supplied by Dr. Fry (Cornell University).
              Leaflets and petioles were isolated from 8 week old
              greenhouse grown plants. The plants were watered and
              fertilized freely. The tissue was immediately frozen in
              liquid nitrogen."
BASE COUNT   139 a 136 c 115 g 180 t
ORIGIN
Query Match      7.7%; Score 116.4; DB 142; Length 570;
Best Local Similarity 52.5%; Pred. No. 1.2e-17;
Matches 289; Conservative 0; Mismatches 246; Indels 15; Gaps 1;
Qy 5 CATTCTCGAATCGCCGAATATCGGCCCTCGGGACCATCGGCCATCGCTCGTTATC 64
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 32 CTTGATAGACAAATGTCAAGTTGCCACACTTCGGCGGCGCAGGGAGGTGACACTCCC 91
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
Qy 65 TCTTACTTTCTTTCGACATTAATCTTGGCTACTCTTCCCTCCGGTCCACCATCTTTCTTCTA 124
    ||||| ||| | | | | | | | | | | | | | | | | | | | | | | | | | |
```

Db 92 TCTTACTTATTTTGTATCATGTTGGTTAGGGTTCCACGGTATACGGCGGATATATTCTA 151
QY 125 TGACTTTCCACATCTTAATCCCATTTTCATGGACACTATATGTTCCAGGCTAAACAATC 184
Db 152 CAAGCTCCCATTTCCAAACCCGATTTTCGTTCAAAACATATATTCCTCTTAAATAATC 211
QY 185 TTTATCGGTCACCTCTTCAACATTTTTCCTCCGTTTGTAGTAAATTTGATTGTTATTCCTAA 244
Db 212 ACTCTCCCTCACTCTCAACACACTATACGCCCTTAGCTGGAAAGTTGCTCGTCCACTAGA 271
QY 245 CACTGATGGTTCGGGTTTTAATAAANAACAGAGAAATAAACAACGTTGAAGGTGATTCGT 304
Db 272 TACAAACGG-----ATATCCTGAGTTACATTTATGTGCACAGGAGATTCGT 316
QY 305 TGTGGTACTTTTGCAGATGTTGTCTTGACTTTTAATAATTTTGCACAGGAATCATCTCG 364
Db 317 ATCTGTTACTTTTTCGAGACTGATATGATTTCAATTTATCTATGTTGACCATCCGG 376
QY 365 AAAATGTGAACATTTTATCCACTTGTACCTTTCATTTGGGAAATGCAATCAAAATATGTGA 424
Db 377 TAATGCTAAGGATTTTATCACTTTGTTCTTAGCTTAGGGAACTTAAGATCACCCGG 436
QY 425 TTGCGTCACGGTCCCACTTTTTCACCTCAAGTGACGTTTTTTCGGGGCTCGGGTATATC 484
Db 437 GGTCCAATAGCTCCGGTCTTAGCCATTTCAAGTGACACTTTTTTCCGGAATCTTGGTGTATC 496
QY 485 ACTAGGAATGACCAATCATATAGCCTTGGTGACGCTAGCACCGGTTCAACTTTTTTGA 544
Db 497 CATTTGGTTTCACTAACCATCATCTTTGTTGGTATGGAGCTACCATAGTGGGTTCATTAG 556
QY 545 AGGGTGGACT 554
Db 557 GGCGTGGCT 566

RESULT 4
BE187619
LOCUS EST336180 KV0 Medicago truncatula cDNA clone pkV0-1613, mRNA 22-JUN-2000
DEFINITION
sequence.
ACCESSION BE187619
VERSION BE187619.1 GI:8666803
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 788)
AUTHORS VandenBosch,K., Endre,G., Hur,J., Moore,J., Beremand,P., Ellis,L.,
Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S., Holt,I.E. and
Fraser,C.M.
ESTs from uninoculated seedling roots of Medicago truncatula
Unpublished (1999)
Contact: VandenBosch K
Department of Biology
Texas A&M University
College Station, TX 77843-3258, USA
Tel: 409 845 7707
Fax: 409 845 2891
Email: kate@mail.bio.tamu.edu
Texas A&M University name:T260795e
TIGR sequence name:MTGAU50TK
More information is available at:
http://chryslie.tamu.edu/medicago
Seq primer: Skmod (CTA gAA CTA gtg gAT CC).
Location/Qualifiers
1. .788
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pkV0-1613"
/clone_lib="KV0"

/tissue_type="Seedling roots"
/dev_stage="Immediately prior to inoculation with
Sinorhizobium meliloti (0 hour)"
/lab_host="E.coli strain XL0R"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA. The cDNA
was directionally ligated into the Unizap XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-Zap phage using Ex-assist
helper phage and propagated in XL0R cells."
BASE COUNT 237 a 188 c 100 g 262 t 1 others
ORIGIN
Query Match 7 3%; Score 110.2; DB 164; Length 788;
Best Local Similarity 54.4%; Pred. No. 3.9e-16;
Matches 279; Conservative 0; Mismatches 213; Indels 21; Gaps 2;
QY 57 TCCTTATCTCTTACTTCTCGACATTACTTGGCTACTCTTCCCTCGGTCACCATCTT 116
Db 5 TCCTCCCCCTAACATCTTTGACATTCCTTGGCTTCTTCTCTCAACCCAACTA 64
QY 117 TTCTTCTATGACTTTTCCACATTTCTAAATCCCATTTTCATGACACTATTTGTCCAGGCTA 176
Db 65 TTTTCTATGAATTTCTCTCATTCATTTCCCATTTTCCACACCACTTGTCCCAAACTT 124
QY 177 AAACAATCTTTATCGGTCACTCTTCAACATTTTTCCTCGTTTCTAGTAAATTTGATTGTA 236
Db 125 AAACAATCTCTTTCTCTCACACTTCAACATTTCTCCCATTTTCCGGTACATTT----- 177
QY 237 TTTCTTAACACTGATGTTTCGGGTTTTTAATAAAAAAACAGAAATAAAACACGTTGAAGGT 296
Db 178 -----CGTTCCTCGCTTGATCTAACCGACCTCACTCGNATTTACTCTTAAC 226
QY 297 GATTCGTGTTGGTTACTTTTTCGAGAATGTTGTCTTGACTTTTAATAATTTGACAGGAAT 356
Db 227 AACTCGGTTTTCATTCACGGTCCGAGAGTCAAAATAGGCACTTTTGAACACTTGTGCTCCGAT 286
QY 357 CATCTTCGAAATGTGAAACTTTTATCCACTTTGTACCTTCATTTGGGAATGCAATCAA 416
Db 287 TACTCAAGAGATGTTAATGAATTTCCACCTTTTGGTACCAAAAGTTGCAACA---AATCTTT 343
QY 417 TTATGTGATTGCTCAGCGTCCCACTTTTTCACCTCAAGTGACGTTTTCCTCGGGCTCG 476
Db 344 TCATTTGAAGTAAAGAGTTCCCTTTGTTAGCTATTCAATAFACATCATTTCCCAATAT 403
QY 477 GGATATCACTAGGAATGACGAATCATCATAGCCTTTGGTGACGCTAGCACGCGGTTCAAC 536
Db 404 GGCTTTTCCATTGGCCTAGCTTTTTCATCATGTTGTTGCTGATGGAAGAACCTTCCACAAT 463
QY 537 TTTTGTAAAGGGTGGACTTCGATTATTCATCT 569
Db 464 TTCATCAAAACATGGTCTTTCATATTTGTTCTTCT 496
RESULT 5
BE920468
LOCUS EST424237 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION cSTB6C8 5' sequence, mRNA sequence.
ACCESSION BE920468
VERSION BE920468.1 GI:10416544
KEYWORDS EST.
SOURCE Solanum tuberosum
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 575)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougrl,O., Buell,C.R., Ronning
C.M., Fry,W.E., Tanksley,S.D. and Baker,B.

```

TITLE      Generation of ESTs from potato leaves and petioles
JOURNAL    Unpublished (2000)
COMMENT    Contact: Cathy Ronning
           The Institute for Genomic Research
           For clone request: please contact Research Genetics, Libraries
           Division tel 1-800-711-6195, email cdna@resgen.com.
           Location/Qualifiers
FEATURES   1. 575
            /organism="Solanum tuberosum"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone_lib="CSFB6C8"
            /clone_lib="potato leaves and petioles"
            /tissue_type="leaflets and petioles"
            /dev_stage="8 weeks old plants"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Tissue was supplied by Dr. Fry (Cornell University).
            Leaflets and petioles were isolated from 8 week old
            greenhouse grown plants. The plants were watered and
            fertilized freely. The tissue was immediately frozen in
            liquid nitrogen."
BASE COUNT 167 a 77 c 139 g 192 t
ORIGIN
Query Match      7.1%; Score 107.6; DB 142; Length 575;
Best Local Similarity 53.8%; Pred. No. 1.6e-15;
Matches 299; Conservative 0; Mismatches 239; Indels 18; Gaps 3;
Qy 805 CATCTTTTACGGTAACCTTGTGGTTATATATGAGTTGCATAGCGAAATCACTCGCTAAAAA 864
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 17 CATCTTTTACTGTACGAGTGTCTTATGTATGAGCTTGCTGATATAATCAAAAGGACGGA 76
Qy 865 TAGGAGAAGAAGGCGCAACGAGTTAGACAG-----TTCAATATCACCATTGATT 918
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 77 TTGGAGAAGAATGATCATAGATAGTAGTGTAAATGGAGTCTCTTTGGATGTGCTGGAGATT 136
Qy 919 GTCGATCTGCTTCATCCACCAATTCACACGCTACTTTGGTAACTGTGTCGACCAT 978
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 137 TTAGACGAGATTCAATCCACACTTCTCAATCTTATTTTGGGAATGTGTAGTTGGGT 196
Qy 979 GTGTCGCCGACCTTAAAAA---ATGTCGTTTGTACTACGGAATAATGGGATGCACCTTGGTG 1035
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 197 GTGTACAGAAGATCAATAAGGCGATGTGACTAGTTGGAAGGAAGGTTTTCGAAATTCGGG 256
Qy 1036 CTAAAGTAATTGGAGCTCTATATCCAAATGATATATAAAGGACGGAATCTTGAAG 1095
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 257 TGGAAATTAATTGGAGAGTCAATCAAAAAAATGAAGGATGAAGATGGGTCTCTTAATG 316
Qy 1096 ATGCCGCGAGATGGCATGAACCTTTTCATGATCCCGGCTAGGAAGATTGGTGTGCTGTA 1155
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 317 GTGATTGTTAAAGAGTTGGCAACGTFAGACCTAATTCGATTTTTCATATGCTGGAT 376
Qy 1156 CACCTAAGCTCAACTGTACGACTTTGATTTGGGTGGGGAAGCGCAATAAGTATGAGA 1215
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 377 GCCCAAAACATGACTTCTATCTGCTGATTTTGGATGGGGAAGGCCGCTGAAGTTAGAA 436
Qy 1216 CTGTTTCAATGACATATAAT-----ACGTCGATTTCTATAAATCAAGCAAAACAT 1266
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 437 TCATTTCTTGACAAATGATGATGGTGAATTTTCGATGCTCTTAGTAAATCTAAGAAAT 496
Qy 1267 CAGCAAGAGATCTTGAAATTTGGATTGAGTCTACCGAGTATGCAAAATGAGGCGTTTCTA 1326
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 497 TTGAGGAGATTAGAGATTGGTTGTCTTGTCTATTACTCAAAATGAATGCTTTTGTCTG 556
Qy 1327 GCATCTTTTGAAGAG 1342
      ||||| ||||| |||||
Db 557 CTATATTCACACACGG 572
      ||||| ||||| |||||
RESULT 6
AW616206
```

```

LOCUS      AW616206      487 bp      mRNA      EST      24-MAR-2000
DEFINITION EST307245 L. hirsutum trichome, Cornell University Lycopersicon
            hirsutum cDNA clone cLHT1D15 5', mRNA sequence.
ACCESSION  AW616206
VERSION     AW616206.1 GI:7322240
KEYWORDS    EST.
SOURCE      Lycopersicon hirsutum.
ORGANISM    Lycopersicon hirsutum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
            Lycopersicon.
REFERENCE   1 (bases 1 to 487)
AUTHORS    van der Hoeven,R.S., Bezzeredes,J.L., Matern,A.L., Holt,I.E., Liang
            ,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
            Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.
            Generation of ESTs from wild tomato (Lycopersicon hirsutum)
            trichomes
JOURNAL     Unpublished (2000)
COMMENT     Contact: David Frisch
            Clemson University Genomics Institute
            Clemson University
            100 Jordan Hall, Clemson, SC 29634, USA
            Tel: 864 656 4366
            Fax: 864 656 4293
            Email: dfrisch@CLEMSON.EDU
FEATURES    5 prime sequence.
            Location/Qualifiers
            source
            1..487
            /organism="Lycopersicon hirsutum"
            /db_xref="taxon:62890"
            /clone="cLHT1D15"
            /clone_lib="L. hirsutum trichome, Cornell University"
            /tissue_type="trichome"
            /dev_stage="mixed stages"
            /note="Leaves of various stages were shaken in liquid
            nitrogen, shearing off trichomes. This procedure yielded a
            mixture of cells which is highly enriched for trichome
            likely with minor contaminations of other types of leaf
            cells."
BASE COUNT 116 a 117 c 91 g 163 t
ORIGIN
Query Match      7.1%; Score 107.2; DB 118; Length 487;
Best Local Similarity 54.4%; Pred. No. 1.9e-15;
Matches 248; Conservative 0; Mismatches 193; Indels 15; Gaps 1;
Qy 64 CTCCTACTTTCTTCGACATTACTTGGTACTCTTCCCTCCGGTCCACCATCTTTTCTTCT 123
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 36 CTCCTACTATTTTGACCATGTTTGGTTAGGTTTCACCTTATTAGGGGATATTATTCT 95
Qy 124 ATGACTTTCACATCTAAATCCCATTTTCATGGACACTATTGTTCCCGAGTAAACAAAT 183
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 96 ACAAGCTCTCCATTTCCAAATCCGATTTTCGTTCAAAACATTATTCCCTCTTAAAAAT 155
Qy 184 CTTTATCGGTCACTCTTCAACATTTTTTCCCGTTTGTGTAGTAATTTGATTTGATTTCCCTA 243
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 156 CACTTTCCTCCTCACTCAACACTATATGCCCTTAGCCGGAACGTTGCTTGTCCACTAG 215
Qy 244 ACATGATGTTTCCGGTTTTTAATAAAACCAAGAAATAAACACAGTTGAAGGTGATTCG 303
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 216 ATCAAAACGGAT-----ATCCTGAGTTACGTTATGTATGACAGAGATTCTG 260
Qy 304 TTGTGGTTACTTTTTCGAGAATGTTGCTTGTGACTTTAATAATTTGACAGGAATCATCTC 363
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 261 TGTCTGTTAAATTTTCTGAGACTGATATGGATTTCAATTTATCTCATTTGGCCACCATCCG 320
Qy 364 GAAATGTGAAACATTTTATCCACTTTGTACCTTGTACCTTATGGGAAATGCAATCAAAATATGTG 423
      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 321 GTAATGCTAAGGATTTTATCATCTTTGTTCTTAAGTAGGGAACCTAAGGATGCACCCG 380
Qy 424 ATTCGCTACCGTCCCACCTTTTTCACCTCAAGTGACGTTTTTTTCCGGGCTCGGGTATAT 483
```

```

Db 381 GGTCCAATTAGCCCGCTTAGCCATTCAAGTGACACTTTTTCGGAATCTTGGGTAT 440
QY 484 CACTAGCAATGACCAATCATCATGACCTTGGTGACG 519
Db 441 CCATTGGTTTCACTAACCATCATGTTGTTGGTGATG 476

RESULT 7
AW221049 606 bp mRNA EST 07-DEC-1999
LOCUS EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF3F1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 606)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Holt, I.E., Liang, F.,
Upton, J., Hansen, T., Craven, M.B., Bowman, C.L., Ahn, S., Ronning, C.M.,
Fraser, C.M., Martin, G.B., Tanksley, S.D. and Giovannoni, J.
TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.
FEATURES
source
location/Qualifiers
1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F1"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the 1cm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN
Query Match 7.1%; Score 107; DB 113; Length 606;
Best Local Similarity 50.4%; Pred. No. 2.2e-15;
Matches 300; Conservative 0; Mismatches 280; Indels 15; Gaps 1;

QY 3 AACATTCTCGAATATCGGCCCTCGGGCACCATCGGCATCGCTCGTTA 62
Db 10 ACCGTATGTGAGCAATGTCAAGTTGGCCACTCCCGGGCCGACGAGGTGATCTC 69
QY 63 TCTCTTACTTCTTCGACATTAAGTGGCTACTTCCCTCCGCTCCACCATCTTTTCTTC 122
Db 70 CCTCTTACTTATTTCGACATGTTTGGTTAGGTTTCGCGGTATGAGCGGATATTATT 129
QY 123 TATGACTTTCACATTTCAATCCCATTTTCATCGACACTATTGTTCCAGGCTTAAACAA 182
Db 130 TACAAGCTCTCCACTTTTCAAAACCCGATTTGGTTCAAAACATTTATCTCTCTTAAAAAT 189
QY 183 TCTTTATCGGTCACCTCTTCAACATTTTTTCCCGCTTTGCTAGTAATTTGATGTTATTCCT 242

```

```

Db 190 TCACCTCTCCCTCACCTCTCAACACATATACGCCCTTAGCGGAAACGTTGCTTGTCCACTA 249
QY 243 AACACTGATGGTTTCGGGTTTAAATAAAAAACCAAGAAATAAAACACGTTTGAAGTGATTC 302
Db 250 GATCAAAACGGAT-----ATCCTGAGTTACGTTTATGTGACAGAGATTC 294
QY 303 GTTGTGTTTACTTTTGCAGAAATGTTGCTTGTGACTTTTAAATAATTTGACAGGAAATCATCCT 362
Db 295 GTGCTGTGTACTTTTTCAGAGACTGATGAATTTCAATATATCTCATTTGGTGACCATCCG 354
QY 363 CGAAATATGCAAAACCTTTTATCCACTTTGACCTTCAATGCGAAATGCAATCAAAATTTATGT 422
Db 355 CGTAAGGCTAAGGATTTTATCATCACTTTTGTCTTAAGTTAGGGAACCTAAGGATGCACCG 414
QY 423 GATTGCTCACGGTCCCACTTTTTCACCTTCAAGTCAAGTCAAGTTCCTCCGGGCTCGGTATA 482
Db 415 GGGGTCCAACTAGCCCGCTCTTTAGCCATTCAGGTGACACATTTTTCGGAATCTTGGTGA 474
QY 483 TCACTAGGAATGACGAATCATATAGCCTTGGTGACCTAGCACGCGTTTCAACTTTTGTG 542
Db 475 TCATTTGGTTTCACTAACCATCATGTTGTTGGTGAGGAGCTACTATAGCAGGGTTCAATT 534
QY 543 AAAGGTGGAATTCGATTAATTCATCTGTTGTTAGATCGGTCCTTTTAAACGAAAG 597
Db 535 AAGGCTGGGCTCTACTCCCAAAATTCGGTGGACATGAACAATCTTATCGAATG 589

RESULT 8
BG097054 508 bp mRNA EST 29-JAN-2001
LOCUS EST461573 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION GSTB45P9 5' sequence, mRNA sequence.
ACCESSION BG097054
VERSION BG097054.1 GI:12587089
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 508)
AUTHORS van der Hoeven, R.S., Bezzerides, J., Holt, I.E., Liang, F., Cho, J.,
Utterback, T., Hansen, C.L., Doan, B., Bougri, O., Buell, C.R., Ronning,
C.M., Fry, W.E., Tanksley, S.D. and Baker, B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info, please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
source
location/Qualifiers
1..508
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB45P9"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 127 a 130 c 87 g 164 t
ORIGIN
Query Match 7.0%; Score 105.8; DB 173; Length 508;
Best Local Similarity 53.7%; Pred. No. 4.2e-15;

```

```
Matches 252; Conservative 0; Mismatches 202; Indels 15; Gaps 1;
Qy 58 COTTATCTTACTTTCTTCGACATTAATGCTTACTCTTCCCTCCGGTCCACCATCTTT 117
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 118 TCTTCTATGACTTTCCACATCTTAATCCCATTTTCATGGACACTATGTTCCAGGCTAA 177
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 178 RACATCTTTATCGGCTACTCTTCAACATTTTTCCTCGTGTGCTAGTATGATGAT 237
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 238 TTCCTAACACTGATGTTCCGGGTTTTTAATAAAAAACCAAGAAATAAACACGTTGAAGGTG 297
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 298 ATTCGTGTTGTTACTTTTTCGAGAAATGTTGCTTGACTTTTAATAATTTGACAGGAATC 357
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 358 ATCCCTCGAAATGTCGAAACATTTTATCCACTTGTACCTTCAATGCGAAATGCAATCAAT 417
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 417 TATGTGATTGGCTACGGTCCGACATTTTTCCTCAAGTGACGTTTTCCTCGGCTCGG 477
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 478 GTATATCACTAGGAATGACGAATCATATAGCTTGGTGGACGCTAGCAC 526
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 526 GCATATCGGTTGGTTTGTAGTAACCATCATGTCGCTTGTGTGGAATATAC 504
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
BE922746 574 bp mRNA EST 02-OCT-2000
LOCUS EST426515 potato leaves and petioles Solanum tuberosum cDNA clone
DEFINITION cSPB21L12 5' sequence, mRNA sequence.
ACCESSION BE922746
VERSION BE922746.1 GI:10448822
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
REFERENCE 1
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
Source
1..574
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSPB21L12"
/tissue_type="leaflets and petioles"
/lab_stage="8 weeks old plants"
/lab_host="SOLR"
note="Vector: pBluescript SK(-); Site1: EcoRI; Site2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. the tissue was immediately frozen in
```

```
BASE COUNT 162 a 123 c 106 g 183 t
ORIGIN
Query Match 6.9%; Score 103.4; DB 142; Length 574;
Best Local Similarity 52.5%; Pred. No. 1.6e-14;
Matches 286; Conservative 0; Mismatches 241; Indels 18; Gaps 2;
Qy 115 TTTTCTCTATGACTTTCCACATTTCTAAATCCATTTTCATGGACACTATGTTCCAGGC 174
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 23 TTTTATCTACAGCTACCAATTTCCAACTCGATTTCTGTTCAAAACCATATATCTCTACTC 82
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 175 TAAACATCTTTATCGGCTCACTTTCACACATTTTTCCTCCGTTTCTAGTATTTGATTG 234
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 83 TTAACATCTACTCTCCCTCACTCTCAAAACACTACACACCTTAGCTGCAACATGCTT 142
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 235 TATTTCTTAACACTGATGTTCCGGGTTTTTAATAAAAAACCAAGAAATAAACACGTTGAAG 294
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 143 GTCCACTAAATTCGAGTGGTT-----ATCCTGAGTTGGGTTATGTGACTG 187
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 295 GTGATTCGTTGTTGCTTACTTTTGCAGAAATGTTGCTTTGACTTTTAATAATTTGACAGAA 354
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 188 GAGATCTATATCTGTTACTTTTATGAGACTGATGATGAATTTCAATCATCTCATTTGGTA 247
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 355 ATCATCTCGAAATGTCGAAACATTTTATCCACTTGTACCTTCAATGGGAAATGCAATCA 414
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 248 ACCATCTCGAAATGCTAAGGATTTTATTCCTCAAT--CCTCAATGGCACAACTAAGG 304
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 415 AATATGATGATGCTGTCACGGTCCGACATTTTTCCTCAAGTGACGTTTTCCTCGGCT 474
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 305 ATGACCCGGGGTCAAACTAGTCCGGCTTTAGGCCATTTCAAGTGACACTTTTTCGGAATC 364
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 475 CGGTATATCACTAGGAATGACGAATCATATAGCTTGGTGACGCTAGCAGCGGTTCA 534
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 365 TTGCATATCGTTGGTTTGTAGTAACCATCATGTCGCTTGTGATGGAATACCATCGTGA 424
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 535 ACTTTTGAAGGGTGAGCTTCGATTTATCAATTCGTTGCTAGATCGGCTTTTTCCTAAGA 594
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 425 AATTCATAAGAACATGGGTTTACTCAACAAATTCGGCGGTGATGAACAGTGTCTTAGAGA 484
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 595 AAGNATCTCCACCGGTTTTTGTATAGATGATTAACATCCACATTTAGATGAATAATAGT 654
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 485 ATGAGTTCAATTCATTTATGATAGTCCGTCGATAGAAAGACCCCTTATGAACAAGGGACGA 544
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 655 TGAGA 659
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Qy 545 TTATA 549
Db ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 10
AW221050 583 bp mRNA EST 07-DEC-1999
LOCUS EST297519 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF3f3, mRNA sequence.
ACCESSION AW221050
VERSION AW221050.1 GI:6532734
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
REFERENCE 1
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 583)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Prisch
Clemson University Genomics Institute
Clemson University
```

100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES

Location/Qualifiers
1..583

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF33"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"
BASE COUNT 142 a 140 c 117 g 184 t

Query Match 6.8%; Score 103.2; DB 113; Length 583;
Best Local Similarity 50.2%; Pred. No. 1.8e-14;
Matches 295; Conservative 0; Mismatches 278; Indels 15; Gaps 1;

QY 3 ACATTCTCGAACATGCCGAATATCGGCCCTCGGGCACCATCGGCATCGTCGTGA 62
DB 10 ACCGTGATTGAGCAATGTCAAGTTGGCCACCTCGCGGGGCAACGAGGTGATATC 69
QY 63 TCTCTTACTTTCTCGACATTACTGGGTACTCTCCCTCGGTCCACCATCTTTCTTC 122
DB 70 CCTCTTACTTATTTGACCATGTTGGTTAGGGTTTCGCCGTATGAGCGGATATTTAT 129
QY 123 TATGACTTTCCACATCTTAATCCCATTCATGGACACTATTGTTCCCAAGGTAAACAA 182
DB 130 TACAAGCTCTCCATTTTCAACCCGATTTTCGTTCAAAACATTTATCTCTCTTAAAT 189
QY 183 TCTTTATCCGTCACATCTTCAACATTTTTCGCTTTCGTTAGTAAATTTGATTTCT 242
DB 190 TCACCTCCCTCACCTCAACACTATACGCCCTTAGCCGGAACGTTGCTTGTCCACTA 249
QY 243 AACACTGATGTTTCGGTTTAAATAAACACAGAAATAAACACGTTCAAGGTGATTC 302
DB 250 GATCAAAACGGAT-----ATCCTGAGTTACGTTATGTGACAGGAGATTCCT 294
QY 303 GTTGTGGTTACTTTGACGAATTTGTCTGTGACTTTAATAATTTGACAGGAATATCATCT 362
DB 295 GTGCTGTGTTTTCGAGACTGATGAAATTCATATCTCATTTGTTGACCATCCG 354
QY 363 CGAAATGTGAAACTTTTATCCACTGTGACCTTCATTGGGAAATGCAATCAAAATATGT 422
DB 355 CCGTAAGGCTAAGGATTTTATACATTTGTTCTCAAGTTAGGGGAACCTAAGGATGCAACG 414
QY 423 GATTGGCTCACGGTCCCACTTTTTCACATTTCAAGTGACGTTTTTTCGGGGCTCGGGTATA 482
DB 415 GGGGTCCAACTAGCCCGCTCTTAGCCATTCAGGTGACACTTTTTCGGAATCTTGGTGTA 474
QY 483 TCATAGGAATGACGAATCATCATCCCTTGTGTGAGCTTAGCAGCGGTTCAACTTTTGTG 542
DB 475 TCCATGTTTTCACTAACCATCATGTTGTTGGTATGGAACACTACTATACATGTTTCAT 534
QY 543 AAGGGTGGAATTCGATTTATCAATCTGGGTAGATCGGTCCTTTTATA 590
DB 535 AAGCGTGGGCTCTACTCCACAATAATACCTGGACATGAACCAATCTTA 582

RESULT 11
BE434257 591 bp mRNA EST 24-JUL-2000
LOCUS

DEFINITION

EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA clone CLEG15H6, mRNA sequence.

ACCESSION

BE434257

VERSION

BE434257.1 GI:9432100

KEYWORDS

EST.

SOURCE

tomato.

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

REFERENCE

1 (Bases 1 to 591)

AUTHORS

Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E., Liang,P., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D.

TITLE

Generation of ESTs from tomato fruit tissue, breaker stage

JOURNAL

Unpublished (2000)

COMMENT

Contact: David Frisch

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366

Fax: 864 656 4293

Email: dfrisch@CLEMSON.EDU

5 prime sequence.

FEATURES

Location/Qualifiers

1..591

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CLEG15H6"

/clone_lib="tomato breaker fruit, TIGR"

/tissue_type="pericarp"

/dev_stage="breaker"

/lab_host="SOLR"

/note="Vector: pBluescriptSKmCuadapt; Site_1: EcoRI; Site_2: XhoI; Fruit were harvested at the breaker stage (first sign of lycopene accumulation on the blossom end of the fruit). Fruit were cut in half and the seeds and locules were discarded prior to freezing the pericarp."

BASE COUNT 159 a 129 c 121 g 182 t

ORIGIN

Query Match 6.6%; Score 99; DB 167; Length 591;

Best Local Similarity 51.7%; Pred. No. 1.8e-13;

Matches 262; Conservative 0; Mismatches 230; Indels 15; Gaps 1;

QY 115 TTTTCTCTATGACTTTCACACTTCCACATCTTAATCCCATTTTCATGGACACTATTGTTCCCGC 174

DB 17 TATATTTTACAGCTCTCCNTTTTCAACCCGATTCGTTCAAAACATATTATTCCTCCCTC 76

QY 175 TAAACAATCTTTATCGGTCACTCTTTCAACATTTTTTCCCGTTTGTAGTAATTTGATTG 234

DB 77 TTAATAATTCACCTCTCCCTCACCTCTCAACACTATACGCCCTTAGCGGAAACGTTGCTT 136

QY 235 TATTTCTCAACTGATGTTTCGGGTTTAAATAAAACACAGAAATAAACACAGTTGAAG 294

DB 137 GTCCACTAGATACAAACGGAT-----ATCCTGAGTTACGTTATGTGACAG 181

QY 295 GTGATCTCTGTGGTTACTTTTGCAGAAATGTTGCTTGTACTTTAATAATTTTACACAGGAA 354

DB 182 GAGATCTCTGTCTGTACTTTTTCGACACTGATGATTAATTTCAATATCTCATTTGGTG 241

QY 355 ATCATCTCGAAAATGTGAAAACCTTTTATCCACTTTGTACCTTCAATTTGGGAAATGCAATCA 414

DB 242 ACCATCCGCTAAGGCTAAGGATTTTATACATTTCTTCTTAAGTTAGGGAACCTTAAGG 301

QY 415 AATTATGATTCGTCACGGTCCCACTTTTTCACCTTCAAGTGACGTTTTCCTCGGGCT 474

DB 302 ATGCACGGGGGTCCCAACTAGTCCCGCTCTTAGCCATTCAGGTGACACTTTTTCGGAATC 361


```

QY 475 CGGTATATCACTAGGAATGACGAATCATCATAGCCTTGGTACGCTAGCAGCGGTTCA 534
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 362 TTGGTATCCATTGGTTTCACTAACCATCATGTTGTTGGTGGAGCTACTATAGCAG 421
QY 535 ACTTTTGAAGGGTGGACATTCGATTAATCAATCGTGTGATAGTCGGTCTTTTAAACGA 594
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 422 GGTTCATTAAAGCGGTGGCTCTACTCCACAATTCGGTGGACATGACAATCTTATCGA 481
QY 595 AAGGATCTCCACCGGTTTTCATAGAT 621
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 482 ATGAGCTAATTCACATTTATGATAGT 508

RESULT 12
BE921494
LOCUS BE921494 564 bp mRNA EST 02-OCT-2000
DEFINITION EST425179 potato leaves and petioles Solanum tuberosum cDNA clone
cstrb13b18 5' sequence, mRNA sequence.
ACCESSION BE921494
VERSION BE921494.1 GI:10447486
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 564)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Uterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdm@resgen.com.
FEATURES
source
1. 564
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cstrb13b18"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/notes="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 150 a 139 c 97 g 178 t
ORIGIN

Query Match 6.5%; Score 98; DB 142; Length 564;
Best Local Similarity 53.2%; Pred. No. 3.2e-13;
Matches 265; Conservative 0; Mismatches 215; Indels 18; Gaps 2;

QY 58 CGTATCTCTTACTTTCTTCGACATTAATGCGTACTTTCCTCCCGTCCGATCATCTTT 117
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 74 CGCTCCCTCTAACTTAATTTATCATATATGTTGGTGGTGGTATATGCGTGGATTT 133
QY 118 TCTTCTATGACTTTCCACATCTCAATCCCATTTTCATGGACATTAATGTTCCAGGCTAA 177
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 134 TATTTCTACAAGCTACCAATTTCCAAATCGATTTTCGTTCAACCAATTAATTCCTACTCTTA 193
QY 178 AACAACTTTATCGCTCACTCTCAACATTTTTCCTCCGTTTGGCTAGTAATTTGATGTAT 237
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 194 AACATTCACCTCCCTCACTCTCAACACACTACACACCCCTTAGCTGGCAACATTCCTGTG 253
QY 238 TTCCTAACACTGATGGTTCGGGTTTTTAATAAAAAACCAAGAAATAAACACGTTGAAGGTG 297

```

```

Db 254 CACTAAATTCGAGTGGTT-----ATCTGAGTTCGTTATGCTGACTGGAG 298
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 298 ATTCGTGTTGGTTACTTTTGCACAATGTTGCTCTGACTTTTAATAATTTGACAGAAATC 357
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 299 ATTCATATCTGTTACTTTTATTGAGACTGATATGAATTTCAATCATCTCATTTGGTAACC 358
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 358 ATCTCTGAAATGTGAAACCTTTTATCCACTTGTACCTTTCATTCGGGAAATGCAATCAAT 417
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 359 ATCTCTGAAATGCTAGGATTTTATCCCAT---CCTCAATGGCACACCTAAGGATG 415
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 418 TATGTGATTCGTCACGTCACGTCCTTTCCTTTCACCTTCAAGTGACGTTTTCCTCGGCTCG 477
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 416 CACCGGGGTCAAACTAGTCCCGGCTTAGCCATTCAGTCACACTTTTTCGGAATCTTG 475
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 478 GTATATCATTAGGAATCAGCAATCATCATAGCCTTGGTGACGCTAGCAGCGGTTCAACT 537
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 476 GCATATCCGTTGGTTTATAGTAACCATCATGTCGCTTGTGATGGAATACCATCGTGAAT 535
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 538 TTTTGAAGGGTGGACTT 555
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 536 TCATAAGAACATGGGGTT 553
    | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 13
BF096979
LOCUS BF096979 482 bp mRNA EST 19-OCT-2000
DEFINITION EST400516 tomato nutrient deficient roots Lycopersicon esculentum
cDNA clone cLEW18O20 5' sequence, mRNA sequence.
ACCESSION BF096979
VERSION BF096979.1 GI:10902689
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 482)
AUTHORS van der Hoeven,R.S., Garvin,D.F., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T.S., Ronning,C.M., Craven,M.B., Bowman,C.L.,
Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Giovannoni,J.J.
and Tanksley,S.D.
TITLE Generation of ESTs from tomato nutrient-deficient roots
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
FEATURES
source
1. 482
/organism="Lycopersicon esculentum"
/cultivar="TA492"
/db_xref="taxon:4081"
/clone="cLEW18O20"
/clone_lib="tomato nutrient deficient roots"
/tissue_type="roots"
/dev_stage="5-6 weeks old"
/lab_host="SOLR"
/notes="Vector: pBluescriptSKmCuAdapt; Site_1: 5' EcoRI;
Site_2: 3' XhoI; Roots were harvested from plants grown
under the following deficiencies/stresses: 10 mM Al, Zn,
P, K, Fe.N. mRNA was isolated from individual treatments.
Proportional aliquots of mRNA of each treatment were mixed
and used for library construction."
BASE COUNT 113 a 119 c 93 g 157 t
ORIGIN

Query Match 6.1%; Score 92.6; DB 144; Length 482;

```


Best Local Similarity 52.8%; Pred. No. 6.2e-12;
Matches 234; Conservative 0; Mismatches 194; Indels 15; Gaps 1;

QY 64 CTTCTACTTCTTCGACATCTACTTGCTGCTACTCTCCCTCCGGTCCACCATCTTTTCTTCT 123
Db 50 CTTCTACTTCTTTTGACCATGTTTGGTTAGGGTTTCGCCGTATGAGCGGATATTATTATTT 109
QY 124 ATGACTTTCCACATCTTAATCCCATTTTCATGACACTATTGTTCCAGCGCTAAACAAT 183
Db 110 ACAAGCTCTCCATTTCAACCCGATTTTCGTTCAAAACATATTCTCTCTTAAAAAT 169
QY 184 CTTTATCGGTCTACTCTTCAACATTTTTCCTCCGTTGCTAGTAAATTTGATTTTCTTA 243
Db 170 CACTCTCCCTCACTCTCAAAACACTATACGCCCTTAGCGGAAACGTTGCTGTCCACTAG 229
QY 244 ACATCATGCTGCTGGGTTTTTAATAAAAAACAGAAATAAAACACGTTGAAGGTGATTCTG 303
Db 230 ATACAAACCGAT-----ATCCTGAGTTAGCTTATGACAGGAGATCTG 274
QY 304 TTGTGTTACTTTTTCGAGATTTGTTCTTGACTTTAATAATTTGACAGAAATCATCCTC 363
Db 275 TGTCTGTTACTTTTTCGAGACTGATGATCAATTTCAATTTATCTATGTTGACCATCCG 334
QY 364 GAAATGTGAAACATTTTATCCACTTTTACCTTCAATGCGAAATGCAATCAATTAATGTG 423
Db 335 GTAAGGCTAAGGATTTTATCACTTTGTTCTTAAGTTAGGGGAACCTAAGGATGCACCG 394
QY 424 ATGCGTCAAGCTGCCACTTTTTCACCTTCAAGTGACGTTTTTTCGGGCTCGGGTATAT 483
Db 395 GGGTCCAACTACCGCGCTTTAGGCAATTCAGGTGACACTTTTTCGGAATCTTGGTGTAT 454
QY 484 CACTAGGAATGACGAATCATCAT 506
Db 455 CCATTGGTTTCACTAACCAATCAT 477

RESULT 14
BG581249
LOCUS BG581249 630 bp mRNA EST 11-APR-2001
DEFINITION EST482982 GVN Medicago truncatula cDNA clone pgvn-64023 5' end,
mRNA sequence.
ACCESSION BG581249
VERSION BG581249.1 GI:13596313
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Medicago.
REFERENCE 1 (bases 1 to 630)
AUTHORS Fedorova,M., Plerson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Town
C.D., Van Aken,S., Utterback,T., Cho,J. and Fraser,C.M.
ESTs from one month old nitrogen-fixing root nodules of Medicago
truncatula, 2001
JOURNAL Unpublished (2001)
COMMENT Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St.Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058
Email: vance004@maroon.tc.umn.edu
University of Minnesota name: M382308e TIGR sequence name:
MTCG98TK More information is available at: http://www.medicago.org
Seq primer: SKmod (CTA gAA CTA gTg gAT CC).
Location/Qualifiers
FEATURES
source
1. .630
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pgvn-64023"
/clone_lib="GVN"
/tissue_type="N2-fixing root nodules"

/dev_stage="effective root nodules harvested one month
post inoculation with Sinorhizobium meliloti"
/lab_host="E. coli strain XL0LR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
effective root nodules harvested one month post
inoculation with Sinorhizobium meliloti. The cDNA was
directionally ligated into the Uni-ZAP XR vector from
Stratagene and packaged using Gigapack III Gold packaging
extracts. Plasmids containing cDNA inserts were excised
from the recombinant lambda-ZAP phage using Ex-Assist
helper phage and propagated in XL0LR cells."
BASE COUNT 186 a 140 c 72 g 232 t
ORIGIN

Query Match 6.1%; Score 91.4; DB 155; Length 630;
Best Local Similarity 53.1%; Pred. No. 1.2e-11;
Matches 265; Conservative 0; Mismatches 201; Indels 33; Gaps 2;
QY 57 TCGTTATCTTACTTCTTCGACATTACTTGGCTACTCTTCCTCCGGTCCACCATCTT 116
Db 75 TCATCTCCCTTACTTCTTTGACATACTATGTTAAGATTACCACCTGTTCAAGAAT 134
QY 117 TTCTTCTATGACTTTCACATTTCTAAATCCCATTTTCATGGACACTATTGTTCCAGGCTA 176
Db 135 TTCTTCTATGAGTTTCCACACCAACCTCATTTTCTTCAAAACCTCTCTTCCCAACTC 194
QY 177 AACAATCTTATTCGGTCACTCTTCAACATTTTTCCTCCGTTTCTAGTAAATTTGATTGTA 236
Db 195 AAAAAATCACTTTCCCATTTGCACTTTTCTACTTTTATCTCTTTTGGTTCATCTCATTTGG 254
QY 237 TTTCTTAACACTGATGTTGCGGTTTTTAATAAAACACCAAGAAATAAAACACGTTGAAGGT 296
Db 255 CCTCATGA-----TTCACAAACCAATCATTAATTCATCAAGGA 296
QY 297 GATTCTGTTGTTGTTACTTTTCGAGAAATGTTGCTTTCGACTTTTAATAATTTGACAGGAAAT 356
Db 297 AACACTCTTTCACCTTACCGTAGCTGAATCTCATGCTGATTTCAATCATTTCTCTGGTAAG 356
QY 357 CATCTCGGAAATGTGAAACATTTTATCCACTTGTACCTTCATTTGGGAAATGCAATCAAA 416
Db 357 AATCTTTCTGAAGCTACACAGATTCATGATTTTATACCTAACCTTGAATATTTCATCA----- 410
QY 417 TTATGTGATTCGCTCAGCGTCCACTTTTTCACCTTCAAGTGACGTTTTCCTCCGGCTCG 476
Db 411 -----CATGCCAAGCTCTATTTTGGCTCTCAAGTTACTATTTTCCAAACTAT 461
QY 477 GGTTATCACTAGGAATGACGAATCATCATGAGCTTTGGTGACGCTAGCAGCGGTTCAAC 536
Db 462 GGATTTTCATGTTGTTATTCATCATCATGCTGTTTATAGATGTTAAACTTCAACTTCT 521
QY 537 TTTTTCGAAAGGTGGACTT 555
Db 522 TTTTCAAAATCTTGGGCTT 540

RESULT 15
AW930934
LOCUS AW930934 497 bp mRNA EST 30-MAY-2000
DEFINITION EST356777 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF42M23 5', mRNA sequence.
ACCESSION AW930934
VERSION AW930934.1 GI:8106335
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 497)
AUTHORS Alcalá,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,

Upton, J., Hansen, T., Craven, M. B., Bowman, C. L., Ahn, S., Ronning, C. M., Fraser, C. M., Martin, G. B., Tanksley, S. D. and Giovannoni, J. Generation of ESTs from tomato fruit tissue Unpublished (1999)

**JOURNAL
COMMENT**

JOURNAL

**JOURNAL
COMMENT**

Unpublished (1999)
Contact: David Frisch

Contact: David Frisch
Clemson University Center

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 4366

Tel.: 864 656 4366
Fax: 864 656 4293

Fax: 864 656 4293

5 prime sequence.

Location/Qualifiers

1 497

1. .497

/organism="Lycopersic

/cultivar="TA496"

```
/db xref="taxon:4081"
```

```
/db_xref="taxon:4081"
/cloc="CFE43M23"
```

```
/clone="CLEF42M23"
```

```
/clone_lib="tomato fruit mature green, TAMU"
```

/tissue_type="fr

```

/dev stage="mature green (3-5 days pre-ripening)"
/tissue_type="fruit pericarp"

```

```
/dev_stage="mature green (3-5 days pre-ripening)"
```

```
/lab_host="SOLR"
```

```
/note="Vector: pBlueScript SK(-); Site_1: EcoR1; Site_2:
```

```

/next= vector::push_back( ), size_1: 2001, size_2:
xhol: cLEF - Fruit were tagged at the 1cm stage and

```

XH01; CLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in

harvested 3-5 days prior to ripening. Fruit were cut in

half to verify the seeds were indeed 'immature' and the

Query Match 5.9%; Score 89.2; DB 122; Length 497;
Best Local Similarity 53.1%; Pred. No. 4.1e-11;
Matches 190; Conservative 0; Mismatches 168; Indels 0; Gaps 0;

QV 264 AATAAAAACCCAGAAATAAAACACGTTGAAGGTGATTCTGTTGTGGTACTTTGCAGAA 323

QY 264 AATATAAACAGAAAATAAAACACGTTGAAGGTGATCTCTGTGTGGTTACCTTTGCAGAA 323

Db 68 AACGGATATCCTGAGTTACGTTATGTGACAGGAGATTCTGTGCTCTTACTTTTTCGAG 127

QY 324 TGTGTCTTGACTTTAATAATTGACAGGAAATCATCTCGAAATGTGAAACTTTTAT 383

Db 128 ACTGATATGAATTCAATTATCTCATTTGGTGACCATCCGGTAAGGCTAAGGATTTTAT 187

QY 384 CCACTTGTACCTTCATTGGGAAATGCAATCAAAATTATGTGATTGCGTCACGGTCCCACCTT 443

Db 188 CACTTGTTCCTAAGTTAGGGGAACCTAAGGATGCACCGGGGGTCCAAC TAGCCCCGCTC 247

QV 444 TTTTCACTTCAAGTGACGGTTTTCGGGGCTATATCAGCAATGACCGAATCAT 503

249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
105

1000

2007 2008 2009 2010 2011 2012 2013 2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025 2026 2027 2028 2029 2030 2031 2032 2033 2034 2035 2036 2037 2038 2039 2040 2041 2042 2043 2044 2045 2046 2047 2048 2049 2050 2051 2052 2053 2054 2055 2056 2057 2058 2059 2060 2061 2062 2063 2064 2065 2066 2067 2068 2069 2070 2071 2072 2073 2074 2075 2076 2077 2078 2079 2080 2081 2082 2083 2084 2085 2086 2087 2088 2089 2090 2091 2092 2093 2094 2095 2096 2097 2098 2099 2100 2101 2102 2103 2104 2105 2106 2107 2108 2109 2110 2111 2112 2113 2114 2115 2116 2117 2118 2119 2120 2121 2122 2123 2124 2125 2126 2127 2128 2129 2130 2131 2132 2133 2134 2135 2136 2137 2138 2139 2140 2141 2142 2143 2144 2145 2146 2147 2148 2149 2150 2151 2152 2153 2154 2155 2156 2157 2158 2159 2160 2161 2162 2163 2164 2165 2166 2167 2168 2169 2170 2171 2172 2173 2174 2175 2176 2177 2178 2179 2180 2181 2182 2183 2184 2185 2186 2187 2188 2189 2190 2191 2192 2193 2194 2195 2196 2197 2198 2199 2200 2201 2202 2203 2204 2205 2206 2207 2208 2209 2210 2211 2212 2213 2214 2215 2216 2217 2218 2219 2220 2221 2222 2223 2224 2225 2226 2227 2228 2229 2230 2231 2232 2233 2234 2235 2236 2237 2238 2239 2240 2241 2242 2243 2244 2245 2246 2247 2248 2249 2250 2251 2252 2253 2254 2255 2256 2257 2258 2259 2260 2261 2262 2263 2264 2265 2266 2267 2268 2269 2270 2271 2272 2273 2274 2275 2276 2277 2278 2279 2280 2281 2282 2283 2284 2285 2286 2287 2288 2289 2290 2291 2292 2293 2294 2295 2296 2297 2298 2299 2300 2301 2302 2303 2304 2305 2306 2307 2308 2309 2310 2311 2312 2313 2314 2315 2316 2317 2318 2319 2320 2321 2322 2323 2324 2325 2326 2327 2328 2329 2330 2331 2332 2333 2334 2335 2336 2337 2338 2339 2340 2341 2342 2343 2344 2345 2346 2347 2348 2349 2350 2351 2352 2353 2354 2355 2356 2357 2358 2359 2360 2361 2362 2363 2364 2365 2366 2367 2368 2369 2370 2371 2372 2373 2374 2375 2376 2377 2378 2379 2380 2381 2382 2383 2384 2385 2386 2387 2388 2389 2390 2391 2392 2393 2394 2395 2396 2397 2398 2399 2400 2401 2402 2403 2404 2405 2406 2407 2408 2409 2410 2411 2412 2413 2414 2415 2416 2417 2418 2419 2420 2421 2422 2423 2424 2425 2426 2427 2428 2429 2430 2431 2432 2433 2434 2435 2436 2437 2438 2439 2440 2441 2442 2443 2444 2445 2446 2447 2448 2449 2450 2451 2452 2453 2454 2455 2456 2457 2458 2459 2460 2461 2462 2463 2464 2465 2466 2467 2468 2469 2470 2471 2472 2473 2474 2475 2476 2477 2478 2479 2480 2481 2482 2483 2484 2485 2486 2487 2488 2489 2490 2491 2492 2493 2494 2495 2496 2497 2498 2499 2500 2501 2502 2503 2504 2505 2506 2507 2508 2509 2510 2511 2512 2513 2514 2515 2516 2517 2518 2519 2520 2521 2522 2523 2524 2525 2526 2527 2528 2529 2530 2531 2532 2533 2534 2535 2536 2537 2538 2539 2540 2541 2542 2543 2544 2545 2546 2547 2548 2549 2550 2551 2552 2553 2554 2555 2556 2557 2558 2559 2560 2561 2562 2563 2564 2565 2566 2567 2568 2569 2570 2571 2572 2573 2574 2575 2576 2577 2578 2579 2580 2581 2582 2583 2584 2585 2586 2587 2588 2589 2590 2591 2592 2593 2594 2595 2596 2597 2598 2599 2600 2601 2602 2603 2604 2605 2606 2607 2608 2609 2610 2611 2612 2613 2614 2615 2616 2617 2618 2619 2620 2621 2622 2623 2624 2625 2626 2627 2628 2629 2630 2631 2632 2633 2634 2635 2636 2637 2638 2639 2640 2641 2642 2643 2644 2645 2646 2647 2648 2649 2650 2651 2652 2653 2654 2655 2656 2657 2658 2659 2660 2661 2662 2663 2664 2665 2666 2667 2668 2669 2670 2671 2672 2673 2674 2675 2676 2677 2678 2679 2680 2681 2682 2683 2684 2685 2686 2687 2688 2689 2690 2691 2692 2693 2694 2695 2696 2697 2698 2699 2700 2701 2702 2703 2704 2705 2706 2707 2708 2709 2710 2711 2712 2713 2714 2715 2716 2717 2718 2719 2720 2721 2722 2723 2724 2725 2726 2727 2728 2729 2730 2731 2732 2733 2734 2735 2736 2737 2738 2739 2740 2741 2742 2743 2744 2745 2746 2747 2748 2749 2750 2751 2752 2753 2754 2755 2756 2757 2758 2759 2760 2761 2762 2763 2764 2765 2766 2767 2768 2769 2770 2771 2772 2773 2774 2775 2776 2777 2778 2779 2780 2781 2782 2783 2784 2785 2786 2787 2788 2789 2790 2791 2792 2793 2794 2795 2796 2797 2798 2799 2800 2801 2802 2803 2804 2805 2806 2807 2808 2809 2810 2811 2812 2813 2814 2815 2816 2817 2818 2819 2820 2821 2822 2823 2824 2825

Search completed: November 5, 2001, 18:01:00
Job time: 14900 sec

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:54:14 ; Search time 8904.87 Seconds
(without alignments)
2643.712 Million cell updates/sec

Title: US-08-894-356C-6
Perfect score: 1522
Sequence: 1 NTCACCACTCTCTCGAATC.....TTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 1344157 seqs, 7733874588 residues
Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba1:*
2: gb_ba2:*
3: gb_ba3:*
4: gb_in1:*
5: gb_in2:*
6: gb_in3:*
7: gb_om:*
8: gb_ov:*
9: gb_pat1:*
10: gb_pat2:*
11: gb_ph:*
12: gb_pl1:*
13: gb_pl2:*
14: gb_pl3:*
15: gb_pl4:*
16: em_ba1:*
17: em_ba2:*
18: em_fun:*
19: em_htgo_hum:*
20: em_htgo_inv:*
21: em_htgo_rod:*
22: em_htgo_hum1:*
23: em_htgo_hum2:*
24: em_htgo_hum3:*
25: em_htgo_hum4:*
26: em_htgo_hum5:*
27: em_htgo_hum6:*
28: em_htgo_hum7:*
29: em_htgo_hum8:*
30: em_htg_inv1:*
31: em_htg_inv2:*
32: em_htg_other:*
33: em_htg_rod:*
34: em_hum1:*
35: em_hum2:*
36: em_hum3:*
37: em_hum4:*
38: em_hum5:*
39: em_hum6:*
40: em_hum7:*
41: em_in:*
42: em_om:*
43: em_or:*

44: em_ov:*
45: em_pat:*
46: em_ph:*
47: em_pl:*
48: em_ro:*
49: em_sts:*
50: em_sy:*
51: em_un:*
52: em_vi:*
53: gb_sts1:*
54: gb_sts2:*
55: gb_sts3:*
56: gb_sy:*
57: gb_un:*
58: gb_vil:*
59: gb_vil2:*
60: gb_htg1:*
61: gb_htg2:*
62: gb_htg3:*
63: gb_htg4:*
64: gb_htg5:*
65: gb_htg6:*
66: gb_htg7:*
67: gb_htg8:*
68: gb_htg9:*
69: gb_htg10:*
70: gb_htg11:*
71: gb_htg12:*
72: gb_htg13:*
73: gb_htg14:*
74: gb_htg15:*
75: gb_htg16:*
76: gb_htg17:*
77: gb_htg18:*
78: gb_htg19:*
79: gb_htg20:*
80: gb_htg21:*
81: gb_htg22:*
82: gb_htg23:*
83: gb_htg24:*
84: gb_htg25:*
85: gb_pr1:*
86: gb_pr2:*
87: gb_pr3:*
88: gb_pr4:*
89: gb_pr5:*
90: gb_pr6:*
91: gb_pr7:*
92: gb_pr8:*
93: gb_pr9:*
94: gb_rol:*
95: gb_rol2:*
96: gb_in4:*
97: gb_pr10:*
98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	Score	% Match	Query Length	DB ID	Description	
1	765.4	50.3	1479	10	E12756	E12756 Perilla oci
2	763.6	50.2	1476	12	AB029340	AB029340 Perilla f
3	193	12.7	1679	12	AB010708	AB010708 Gentiana
4	193	12.7	1703	10	E12753	E12753 Gentianatri
5	172.6	11.3	1508	10	E12757	E12757 Senecio cru
6	153.8	10.1	1622	10	E12754	E12754 Gentianatri
7	153.8	10.1	1622	12	AB026494	AB026494 Gentiana
8	106.4	7.0	66237	12	AB016892	AB016892 Arabidops

9 102.2 6.7 101176 12 AC002560 Genomic s
10 87.8 5.8 119914 12 AC003027 Arabidops
11 77.8 5.1 85690 12 AB028618 Arabidops
12 77.8 5.1 136047 74 AC069470 Arabidops
13 71.2 4.7 1648 9 AX025514 Sequence
14 70.2 4.6 1167 53 CDS07360
15 69.6 4.6 82360 13 AP000606 Arabidops
16 69.6 4.6 341064 75 AC074226 Arabidops
17 65.8 4.3 224448 84 PF041425 Plasmodi
18 63 4.1 11829 4 AE001425
19 62.4 4.1 49999 9 AX015915 Sequence
20 61.6 4.0 1741 5 AF305060 Dictyoste
21 61.6 4.0 136047 74 AC069470 Arabidops
22 60.8 4.0 4596 5 AF310896 Dictyoste
23 60.6 4.0 5481 96 PFSRPTG X77854 P.falciparu
24 60.6 4.0 201508 70 AC026290 Homo sapi
25 59.2 3.9 4362 6 DU060170 Dictyostell
26 59.2 3.9 4617 96 SFAJ5398 Dictyoste
27 59.2 3.9 85690 12 AB028618 Arabidops
28 59 3.9 1116 53 CDS07117J AL425357 clone BA0
29 59 3.9 6372 6 DDICNPA M23449 Dictyoseliu
30 59 3.9 77835 84 PF0413P2_3 Continuation (4 of
31 58.8 3.9 2452 5 AF337815 Dictyoste
32 58.8 3.9 83110 84 PF0413PB AL109814 Plasmodi
33 58.8 3.9 156060 60 AC004153 Plasmodi
34 58.8 3.9 173003 68 AC024245 Homo sapi
35 58.4 3.8 1598 6 DMYSOLC X54161 D.discoideu
36 58.4 3.8 2426 96 PF053326 U53326 Plasmodiu
37 58.4 3.8 146882 61 AC010933 Homo sapi
38 58 3.8 2914 5 AF310893 Dictyoste
39 58 3.8 4458 6 DDIDP87 D13973 Dictyostell
40 58 3.8 310779 60 AC005140 Plasmodi
41 57.6 3.8 46013 83 AP001873 Homo sapi
42 57.6 3.8 174640 64 AC017044 Homo sapi
43 57.4 3.8 797 53 CDS07A6G AL436190 J7 end of
44 57.4 3.8 1147 53 CDS073CX AL427351 clone BA0
45 57.4 3.8 6001 5 AF310894 Dictyoste

ALIGNMENTS

RESULT 1
E12756 LOCUS
DEFINITION Perilla ocimoides mRNA for acyltransferase, partial cds.
ACCESSION E12756
VERSION E12756.1 GI:3251588
KEYWORDS JP 1997070290-A/4.
SOURCE unclassified.
ORGANISM unclassified.
REFERENCE 1 (bases 1 to 1479)
AUTHORS Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y., Yonekura, K., Mizutani, M. and Kusumi, T.
TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 4 18-MAR-1997; SENTRY LTD
COMMENT OS Perilla ocimoides
PN JP 1997070290-A/4
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO MASAHITO.
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAKI PC
C12N15/09, A01H1/00, C07H21/04, C07K14/42, C12N9/10, (C12N9/10, PC C12R1:865),
PC (C12N9/10, C12R1:19);
CC strandedness: Double;
CC topology: Linear;
FH key Location/Qualifiers
FH

FT source 1..1479
FT /organism='Perilla ocimoides' FT
FT /tissue_type='leaves'
FT /clone='pSAT208'
FT CDS 1..1343
FT /product='acyltransferase'.
FEATURES
source Location/Qualifiers
1..1479
/organism='unidentified'
/db_xref='taxon:32644'
BASE COUNT 420 a 316 c 331 g 411 t 1 others
ORIGIN
Query Match 50.3%; Score 765.4; DB 10; Length 1479;
Best Local Similarity 71.1%; Pred. No. 2.5e-147;
Matches 1056; Conservative 1; Mismatches 417; Indels 11; Gaps 3;
Qy 8 CCCTCTCGAATCTCTCCGAGTGGCGGCTCCAGGCACGGTGGCTGAGCAGTCACTCC 67
Db 1 CCGTGATCGAAACGTGTAGAGTTGGCGCCGCCGCGACTCGGTGGCGAGCAATCGGTGC 60
Qy 68 CGCTCACCTTCTTGACATGAGTGGCTGCAATTTCCACCCCAATGCTTCAGCTTCTTCT 127
Db 61 CGCTCACATTTCTTGACATGAGTGGCTGCAATTTTCATCCCATGCTTCAGCTTCTTCT 120
Qy 128 ACGAATCTCCCTGTTCCTCCAAACCCGCTTCTCGAACCCTGCTCCGAAACCTCAACAA 187
Db 121 ACGAATCTCCCTGTTCCTCCAAACCAATTTTCGAAATCTATCCGAAACCTCAACAA 180
Qy 188 CTTATCTCTCAACCTCAACACATTTCTCCGCTTCTCATGCAATTAATCTACCTCTAT 247
Db 181 CTCCTCTAAACCTCATACATTTCTCCCTCTCTCATGCAATTAATCTACCTCTAT 240
Qy 248 CGCGGAGAAATCGCGAGTTCGGGTATACAGACGGTGTGCTCGGTTCCTTCAGATTA 307
Db 241 CCCGGAGAAATCGCGAGTTCGGGTATATCGGTATCTATCCGGGAGCTCGGTTCCTTCA 300
Qy 308 TGGAGTCTGTCGGAGATCATCCGATTCGGCTCATAAATCTACTGCTTGGCCCTAGGG 367
Db 301 CAGAATCTAGGAGAGCTTCGATGATCTCGTCCGAAATTCGTCAGAAATCTCCGTTAGGC 360
Qy 368 ACGATTATGAAGATCTCCAGCTGCGCGGATGTCGAGGAAATCTCATCGGAAATTTGTTTC 427
Db 361 TCTACAATTTGTCCCTAAATTCGCGCCCATGTCGAGAAATCCGATAGAAACCTCTCC 420
Qy 428 AAGTTTTCAGCGTCAAGTACTGTTTCCGCTGCGGGGTGTCATCGGAAATTAACGA 487
Db 421 AAGTTTTCGCGTGCAGTGACTCTTTTCCAGGCGGAGCGTTCGATTTGAATAGCAA 480
Qy 488 CGCACACACGTTAGGATGCTCATGCTTTGTAGGGTTTATGAAGATTTGGGCTTCCA 547
Db 481 CGCATCACACGTTAGGAGCTGCTTTTCCGCGGCTTTATACGGCTTGGTCTTCAA 540
Qy 548 TCACCTAATTCGGAGGAGATGATGAATTTCTTGGAGGAAAGGTAATGTTTGGCGGTTT 607
Db 541 TGAGCAACACATTTGAATTAAGA---TGAAGTGAAGATTTAAATCTTTGCCAGTTT 597
Qy 608 TCGACCGATCGCTCGTGAATTTATCGGCTTAATTTGGACACATATTTATGGAAACACGGC 667
Db 598 TCGATAGTCGCTCATAAATATCCGAGAAATTTGATCCATTTATTTGGAGAAACGGC 657
Qy 668 AGAAACGTCGTTGAATTCGAGATCTCTTTACCGACGATCGGATTCGAGTACCT 727
Db 658 TAAATTTCTTTCGATCTCGTCACTCCCTCATTTACCGACGAGCGCATTCGAAACCGT 717
Qy 728 ACCTTTTACCCCAATCTGAAATTAAGAAATTAAGAGGTTTGAATTCAGAGAAAGCCCAA 787
Db 718 TCGTTTTCACCCCAATTCGAAATTAAGAAATTAAGAGGTTTGAATTCAGTCCAGAGTTCAA 777
Qy 788 ATGTAGTTAATCTCTCTCTTCCTTCGCGGATCGGAGCTTATATCTGGACCGGCATCGCA 847
Db 778 GTTAGTCCATCT 837

[illegible]


```

Db 473 CTATTACCACTTTTTCGTTTACAGGTGACGGTGTCTCCACACCGGTATATGCAATTGGA 532
Qy 481 ATAACGACGACACACACCGTTAGCGATGCTCATCGTTTGTAGGTTTATGAAGAGTTGG 540
Db 533 CGCAATCTTCATCAAGTCTTGTGGTACCGAGTCTTTTTCGCAATTTTAAATAAATATGG 592
Qy 541 GCTTCCACTCAATAATTCGGAGGAGATGATGAATTTCTTTGGACGAAAAAGGTGAATGTTTG 600
Db 593 GTTTTGGTTGCAAAATCCAATGGAGATTC---ATTAAAGTTCTTCCACTTCTTCTCTA 649
Qy 601 CCGGTTTTCGACCGATGCTGTAATATATCGCCCTAAATTTGGACACATATTTATFGGAAC 660
Db 650 COTATGTACGACAGATCTGTGTGCAAGATCCATTTTCATATTCGTGCGAAAAATCTACAA 709
Qy 661 -----AACGGCGCAAAACGCTCGGTGGAATC-----GCAGCATCCATCTTTACCGAG 708
Db 710 GAAAGAAACTGCTCAATCTCAGGGCACACTACTGTCTTAAATCCAGCAATTTCTTAA 769
Qy 709 GATCGGATTCGAGCTACCTACCTTTTTCACCAATCTGAAATTAAGAAATTAAGAGGTTTG 768
Db 770 GATGAAGTTTCGAGCCACCTTCATCTACACCTATTGATATCATGAAGCTCAAGAAATTC 829
Qy 769 ATTCAGAGAAAGCCCAAAATGTA-----GTTAATCTCTCTCTCTCTCTCTCTCT 813
Db 830 ATTTCTGTCAAAAAATCGCAACTTAACCGGTAGTAGTAATATATATCTGTCAACTTTCAC 889
Qy 814 GCGATTCGCGAGCTTATATCTGGACCGGCATCGCAACT-----CGTTCGGAGATTAC 864
Db 890 GTGACATCTGCACTGATCTGACATGCTGTGCGAATCATTTAGACACCGCTCGTAGAGAG 949
Qy 865 AAAGACGTGGATGACGACAAACCGCTTTCTTTTAAATTCGGATCGATTTAAGCGCGGT 924
Db 950 AAGGTGGAAGAGGATAACATGCAACAACTTATGTGCTTTCATCAACTCGCGCAACGT 1009
Qy 925 TTGGATCCGCGGCTCCGGGAACTACTTCGGAACCTGCTATATCGTTTTCGATGCGCGAG 984
Db 1010 TTTGCTCCGCGGATACCTTCAAAATTTACTTTGGAATTCGATAGTGCCTTGTATGTTGGGA 1069
Qy 985 ATCCTCGCGCGGATTTGCTCGGAGATGAAGGGTGTTCGGCGAGCTGAGCGCATCGCG 1044
Db 1070 TCGACTCATGAGCACTTTGAGGAATGAGGGTTGTCGGTAGCTGCAACCGCCATCGGA 1129
Qy 1045 GCGGAAATAGAGAGAGGACGAGCGACAAGAA---GATTCTAGAACTGTGAGAACTGG 1101
Db 1130 GATGCTATCCATAGAGGTTTACATGCACTACGAAGAAATCTGAGAGGAGATTGGATATCG 1189
Qy 1102 CCGTCTGAGATTCGGAAGCCTTGCAAACT-----GTTATTTCTCGGTGCGCGGA 1152
Db 1190 CCGCCCGGATCAACATCTCGCGCACCAAGGTCGACGCTCATTTATGCTGTTGGATCCGCA 1249
Qy 1153 TCGAGCAGGCTTGATCTTTACGGCGCGGATTTTGGATGGGTAAGCGGTGAAGCAAGAG 1212
Db 1250 CAACGCATATGTCATGATTTTGTATGACAGATTTTGGTGGGAAAGCTTGAAGAGCATGAA 1309
Qy 1213 ATACTGTGCTGATGAGAGAAAGTTTACGATGTCGTTGTGTAAACCGAGGATGCTGCC 1272
Db 1310 TCTGTTTCAACTTAATCTTCGCAACACTAATTTTGTATCTCTCGTCCAGAAGATTTAAA 1369
Qy 1273 GGAGGATTTGGAGTTGGATTTGCTTTTTCGCAAGGAGGAAATTTGCAAGCTTTTGTATGAT 1332
Db 1370 GGAGCAGTTGAGCTTGGCATTTCTTTTGGCTTGAAGATAGGATGGAGCATTTTGCACCAT 1429
Qy 1333 TTTGCGGAGGGAATAAAGGGT 1353
Db 1430 TTTAGCAATTTTCATCAATAGT 1450

RESULT 7
AB026494 1622 bp mRNA PLN 04-APR-2000
LOCUS AB026494
DEFINITION Gentiana triflora GAT106 mRNA for acyltransferase homolog, complete
cds.
```

```

AB026494
AB026494.1 GI:7415596
acyltransferase homolog.
Gentiana triflora cDNA to mRNA.
Gentiana triflora
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Gentianales; Gentianaceae; Gentiana.
1 (sites)
Yonekura-Sakakibara,K., Tanaka,Y., Fukuchi-Mizutani,M.,
Fujiwara,H., Fukui,Y., Toshihiko,A., Yamaguchi,M. and Kusumi,T.
Molecular cloning and biochemical characterization of
hydroxycinnamoyl-CoA:anthocyanin
3-O-glucoside-6-O-hydroxycinnamoyltransferase from Perilla
frutescens and diverse plant acyltransferase homologs
Unpublished (1999)
2 (bases 1 to 1622)
Tanaka,Y. and Yonekura-Sakakibara,K.
Direct Submission
Submitted (21-APR-1999) to the DDBJ/EMBL/GenBank databases.
Yoshikazu Tanaka, Suntary Ltd., Institute for Fundamental Research;
Wakayama-dai 1-1-1, Shimamoto, Osaka 618-8503, Japan
(E-mail:Yoshikazu.Tanaka@suntary.co.jp, Tel:81-75-962-8807,
Fax:81-75-962-8262)
FEATURES
Location/Qualifiers
1..1622
/organism="Gentiana triflora"
/db_xref="taxon:55190"
35..1474
/gene="GAT106"
35..1474
/gene="GAT106"
/codon_start=1
/product="acyltransferase homolog"
/protein_id="BAA93452.1"
/db_xref="GI:7415597"
/translational="MAGNSEDIKVLEKCRVAPPPDAVEFTVPLSFDDMRWLISDAEH
SLVSLTIAESMDFDYLAGDHORDSYKFNDLIPOLPEPIVTSGDVLPLEALQVTFESN
LHMFYRPHCPNSKFTISSTKSLVLFKHLPLAGNLIPWDSDDRMPELRYKKGD
TGICRNLHVLGVDASSFLHFNKLVLDKSNGLSKFLPLSLPMDRSVVDPEH
IRKNIYERKLLSQGTFTNPAINPAISKDEVRATFILPIDIMKLKFTISSNNRLTGS
SNYLSFTTTSALITWLSKSLDTVVREKVEDKHAANLCAFINCHQREAPPYIPQNY
FGNCIVPCMVGSGTHEQLVNEGLSVAATAIGDAIHKRLHDYEGILRGDWISPPRTSA
APRSTLTYVYVGSQNRVHDFDADFQWGLKLEKHESVSTNPSTLILISRRRFRKGALE
GISLPKRMDFATIFTNFINSLHVRPL"
BASE COUNT 478 a 345 c 322 g 477 t
ORIGIN
Query Match 10.1%; Score 153.8; DB 12; Length 1622;
Best Local Similarity 49.4%; Pred. No. 1.2e-21;
Matches 692; Conservative 0; Mismatches 652; Indels 57; Gaps 9;
Qy 4 ACCACCCCTCTCGAATCCTCCGAGTGGCGCGCTCCAGGCACGGTGGCTGAGCAGTCA 63
Db 56 ATCAAGTCTTGTGAAATGCGGTGTCGCCACACCGGACGCGCGAGTTTACA 115
Qy 64 CTCGCGCTACCTTCTCGACATGACGTGCTGCAATTTCCACCCCATGCTTCAGCTTCTC 123
Db 116 GTCCCACTGCTGTTTTCGACATCGATGGTGTGATCTCTGATCGAGAACACATCTGCAT 175
Qy 124 TTCTACGAACCTCCCTCTGTTCACAAACCCCGCTTCCTCGAAACCGCTTCGCCAAACTCAAA 183
Db 176 TTCTACAGATTCGCGCATCTTGTCCCAACTCTA--AATTATCATTTTCATCATTA 232
Qy 184 CAATCCCTTATCTTAACCCCTCAAAACACTTCTTCCCCCTTTTCATGCAATCTAATCTACCT 243
Db 233 TCGTCCCTTTTCCCTTGTCTCAAAACACTTCTTCCGTTAGCCGGGAATTTGATTGGCCG 292
Qy 244 CTA---TCGCGCGGAGAAATGCCGCGAGTTCGGGTATCAGAACGCTGACTCGGTTCTTTC 300
Db 293 GTAGATTCTCCCGAGAAATGCCGCGAGTTCGGTTTACAAGAAAGGGGACTCCGTTTCTTTA 352
```


JOURNAL

Submitted (15-JUL-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th Street and
Hamilton Walk, Philadelphia, Pennsylvania 19104-6018, USA

REFERENCE
AUTHORS

7 (bases 1 to 101176)
Cheuk,R., Shinn,P., Brooks,S., Buehler,E., Chao,Q.,
Johnson-Hopson,C., Khan,S., Kim,C., Altafi,H., Bel,B., Chin,C.,
Chlou,J., Choi,E., Coni,L., Conway,A., Gonzalez,A., Hansen,N.,
Howing,B., Koo,T., Lam,B., Lee,J., Lenz,C., Li,J., Liu,A., Liu,J.,
Liu,S., Mukharsky,N., Nguyen,M., Palm,C., Pham,P., Sakano,H.,
Schwartz,J., Southwick,A., Thaveri,A., Toriumi,M., Vaysberg,M.,
Yu,G., Davis,R., Federspiel,N., Theologis,A. and Ecker,J.

TITLE

Direct Submission

JOURNAL

Submitted (19-JUL-2000) Arabidopsis thaliana Genome Center,
Department of Biology, University of Pennsylvania, 38th and
Hamilton Walk, Philadelphia, PA 19104-6018, USA

COMMENT
FEATURES

On Jul 15, 2000 this sequence version replaced gi:2618677.
Location/Qualifiers

source

1. 101176
/organism="Arabidopsis thaliana"
/db_xref="taxon:3702"

/chromosome="1"

/clone="F21B7"

join(736..1038,1276..1713,1874..2119,2205..3032,
3117..3274,3483..3610,3725..4134,4313..4435)

/note="unknown protein; similar to ESTs dbj|AV529515.1,
dbj|AV529266.1, dbj|AV54950.1, dbj|AV523761.1,
dbj|AV523761.1, dbj|AI996703.1, and dbj|AV546743.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.1"

/protein_id="AAF86508.1"

/db_xref="GI:9280639"

/translation="MQPLPARSGDEGFNSHPLLVLVAGDETNGTGLGHFSQNGSL
ARGSDSKAGDAINPTTVRYSLRSHSVYVLRFRSSVCMIRCSRVAVGLANGI
YCDALRLKNSVLTVPQVRQGTTRVNGTGPVAGPRLVIAKSKSMKTGR
LPDPTFSPSPSSSGSSGMARYAMESSKOLANGLNGLMGDTKYSKYCQDML
PSTQSPASNAWKGVGSGDAENAGYAVKDLVSGALVSQFAKHTSIKALCFDP
SGTLTAVTCGNNINVTMPGRSHNAPGDLSEWESSHVLFLKLRGITSATVQDI
CFSQSQWVAIISKQCHIFVLNMSGDAAPCEGEETRLPASLPMWFQSLSS
NQSLSPPTAVLSVSRIRKYSFGWLTNTATATAGKVFVPSGAAVAFKSVTH
DLQNSRTNALEHLVTPSGHVQHELLSVCTESPENGLRVOKTSHVQVQDDLRV
KVEPIQWMDRLEETRLPRLPTEKOYDLETVSNHLTSHEDACLSDMNSHFS
EDKYLKSKPPERSHCYLSNFVKTSGMLPWQNSKISFHVMDSPRDSKGTGEF
EIEKPAHEIEIKOKLLPVDFHSTKATLEDRFSMKCYHTSATGSHOVNGKICQDI
INCHSKPGSTESAESSPEGTKOMENLHSDHMSNSIKSLPLYPTVNGIYKTEKN
ANGMKPVTAKLSTLKTETGTTPPILITDSVNEQMLSTGKPMWGFGLHEHC
KAVADPKEEHLKKLDEVTNVLNWNNTKLGQDKMVSQVLAFAFCILYKDFL
RVGNFNSVFFLLYKTLWRK"

complement(4838..6223)

/note="hypothetical protein"

/codon_start=1

/evidence=not_experimental

/product="F21B7.2"

/protein_id="AAF86541.1"

/db_xref="GI:9280672"

/translation="MASCIGELHFTHLHIPVTINQQFLVHPSPPTPANQPHLSLYLS
NLDDIICARVETSVYPTNNRESFVLKRLQDALSEVLVYVPLSGRLREYENGKL
EVFPEQGVLMVYANSMDLADLITVNPALPLIFRNPGEAYKILEMPLLIAQ
VTPECCGFSGLRCLCHICDGGAMQFLGSAWATKGLIADPEVPWDRTEFKPRN
PMWKYPHIEPLTEENSLNMDTKPLQKCYRISKEFCQCRKSTAQGEDPLVCS
TFDAAAHIWRSWKALDKVLDYNLRKLTFSVNRTRLETKLRKSGYNNVCLACAM
SVSEYELNLSKTRTLVQDARLRSYEDYLRSDYVDVVRKPRLEFGKLTITQWTR
FENYETADFGWGPVYAGIDLRPTQCVLLPQGGVSGNQDSMVVCLCLPTTAVHT
FTRLNLDHK"

complement(6700..7357,7482..7884,8620..8839,
8920..9241,9758..10380))

/note="unknown protein; similar to ESTs gb|T04168.1,
gb|T21964.1, and gb|AW004462.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.3"

/protein_id="AAF86540.1"

/db_xref="GI:9280671"

/translation="MGHDSFCYLLIVLCALRCGLIALMOICALQKERRSKMESSDRS
SOAKAFETKGVKLVASGIKEIPAMFHTPDLTSLKOTAPPSOOLITPTVDLKGG
SMDLISRSVVEKIGDAERGWFFQVNVHGISVEVMEKREIRRFHQDEPEVKKREY
SRDTRDLYISNIDLTCHKANWRDTLACYMAPDPKLODLPAVCGEIMMESKOL
MTLGEFLELLSEALGNHLKMDGCAKSHIMFGYPPCPQDPLTLGIGSLDFDFSF
ITILLQDNGIGLQVHDCQWVDPVPCALVINDLLQILISNDKFI SAERHVIANGS
SEPSIMPCFVSTFMKPRIYGPICKELSEQNPAKYRDLTITEFSNTFRSQTIRTE
IPALFRATPATLASKSPPPKHILTPTVDLKGASVVEKIGEAEEKGLFHLVHNP
VEVLERMIQIRGHEQEPKAKRFYSRDTRDVLVYFESNHDQLNSEASWRDITLGCYT
VGPRLLEDPAVCGEIMKESIMSKELFELLSEALGNHLKMDGCAKSHIMFGYPPCPQDPLTLGIGSLDFDFSF
VGOHYPPQDPLTIGINKHTDISFLTLQDNGVGLQVHQEYHIDVTPVPGALVIN
IGDFQLKMLLRSCFISCMIMGLILFYCVQLITNDKFI SAERHVIANGSEPRFT
SVAIVFTFMRAVSRYGPIKDLLSAENPAKYRDLTITEFSNTFRSQTIRFSLDAPKLHHEK
I"

10833..11348

/note="unknown protein; similar to EST emb|Z30724.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.4"

/protein_id="AAF86509.1"

/db_xref="GI:9280640"

/translation="MRGFSVSSHLWTSVEIGEVFLGTSWTLDAVAVAEDAVDMWVS
ACLMVKNKDLRLVLPVAGLSEPSFFSLRFLSPSIRIEGCFAGKFSDLRFRSE
AWKCVVLRGWLQDQIGRWSVGNMVFVRFVAGPASSLVRDLQDSYTSANDLVPVV
LQSSGPGIRRY"

join(11965..12033,12173..12303,12448..12519,12596..12685,
12756..12915,12985..13041)

/note="similar to ATHP1 dbj|BAA37110.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.5"

/protein_id="AAF86510.1"

/db_xref="GI:9280641"

/translation="MNTIVVAQLQRQDVIYLSYQOGLDNQFSELKLODECTPDF
VAEVSLFFDDCSKLNTMSISLERPDNDFKQVDSGHQLKSSSSLCFLVLSGARR
VKNVCISKECCDVQNRREGYSSNHTNDNDINHRCRLCQLQVDYKMLKRLQLDLF
NVSPKLCENLLILLEKQILQAGGTIPQVDIN"

16495..17682

/note="similar to protein kinase pir|JAS57676; similar to
ESTs gb|A1993651.1, dbj|AV538995.1, dbj|AV522524.1, and
dbj|AV538995.1"

/codon_start=1

/evidence=not_experimental

/product="F21B7.6"

/protein_id="AAF86511.1"

/db_xref="GI:9280642"

/translation="MRTKLWCLMFLRLRFGFTTEALDIPVDFLALQAIRKSLDDLPG
SKFFESWDFSDPCGFAGVYNGDKVLSLNLGDPAGSPGLSGRIDPAIGKLSALTEL
SIVPGRIMGALPATISQKDLRLAISRNFISGEISGLSEVRGLRTIDLSYNQLTGT
ISPSIGSLPELSNLIILCHNLTGSIPLPFLSQTTLRIDLRKNSLTGSIASLPSPSLQY
LSLAWNLTGVSVYHLLRNOLNYLDLSLRFTGTIPARVAFPTITNLQLRNPFEGFL
IQPANOVTISTVDLSYNRFSGGISPLSSVENLYLSNRFTGEVPASFVERLLSANIQ
TLXLQHNLEIGIQISPAAEIPVSSSLCQVNCMVPLQTTCPLKAGPQKTRPTTQCTE
WRG"

complement(18462..18515,18612..18742,18840..18975,
19064..19139,19214..19410,19497..19595,19668..19752,
19865..20103,20179..20253,20345..20649,20798..20902,
21088..21216,21356..21437,21513..21608,21780..21849,
21929..22025,22408..22502,22586..22659,22810..22888,
23041..23101,23206..23315,23411..23556))

/note="hypothetical protein"

/codon_start=1

/evidence=not_experimental

/product="F21B7.7"

/protein_id="AAF86539.1"

/db_xref="GI:9280670"

/translation="MPEKFKRRSLNFICLPKLKECARGOILGYKMETKAI GFHLSLP
PQNTVCGHTLTAVFVNSHQILIFGGSTAVANHNSLSLPEISLDGVNSVHSPFLVLR
KWRNLPIGDVPPSRACHAALVYTLIIQGIIGSPGSDGVYMLDMTNKMKFLV
GGTSPRYGHVMDIAAQRALVIFSGNNGNEILDWTALDTRGPFSDRLNPSGNQPS
GRWYASSSREDGIFLCGGIDHSGVTGLDGYGLKMSDNDVWTPVPAVSPRYQHTA
VFGSKLHVIGGLILNRLRDLGEAVADTNPQETSASGANRQNOYQLMRRCHHAAASF

CDS

CDS

CDS

Query Match 6.7%; Score 102.2; DB 12; Length 101176;
Best Local Similarity 45.4%; Pred. No. 7.5e-11;
Matches 620; Conservative 0; Mismatches 718; Indels 27; Gaps 6;

QY 7 ACCCTCTCGAATCTCCGGAGTGGCGGCTCCAGGACACGGTGGCTGAGCAGTCACTC 66
DB 37271 AACATCATCGAGACCTGTGCACATCTCTCCCAAGGACCGCTTGCATCCACCACCTCT 37330
QY 67 CGGCTCACCTCTTCGACATGAGTGGCTGCAATTTCCACCCCATCTTCAGCTTCTCTTC 126
DB 37331 CCTCTCACCTCTTCGATGCCCCCTGGCTCTCTCCCACTCGCGGATTCCTCTCTTC 37390
QY 127 TACGAACCTCCCTGTTCCTCAAAACCCGCTTCTTCGAAACCGTGTCCGAAACTCAAAACAA 186
DB 37391 TTCTCTTACCAAACTCAACTGAATCTTTCCCTCCAGAGCTTTGTACCCCAACCTCAACAT 37450
QY 187 TCCCTTATCTAACCCTCAACACTCTTCCGCCCTTTTCATGCAATCTAATCTACCCCTCTA 246
DB 37451 TCCCTCTCATCACTCTCCCAATTTCTTCCCTTACGCGGCTTAACTGATTTATCCGCCCT 37510
QY 247 TCGCCGGAGAAAAATGCCGGAGTTCGGGTATCAGACGGTGAAGTCTCTCTTCTTCAAGATT 306
DB 37511 CGTCCGGACCTCCATATTTACACTACAACGCTGGGAAGACTCTCTGTTTTCACCGTA 37570
QY 307 ATGGAGTCT-----GTCGGAGATCATCCGCTATCCGCTCATAAATPACTACTGCTTT 357
DB 37571 GCAGAGTCTACTGAAACGACCTTCGACCAACTGAAGTCTGATTCCTTAAAGATATCACT 37630
QY 358 GCCCTAGGAGCATATGAAGATCTCCAGCTGCGGCCGATAGTGAGGAACTCATCGG 417
DB 37631 GTGTTGCATGGCTCTTGCCCAAGTTACCTCTCCTCAGCTCTCTCCGGAAGAAATTCAG 37690
QY 418 AAATGTTTCAAGTTTACCCGCTGCAAGTACTCTGTTTCCCGGTGCGGGGTGTCATC 477
DB 37691 A---TGCGACCAATATGGCAATGAAGTCAACCATCTTCCCTGGAGCTGGAATCTGTATA 37747
QY 478 GGAATAACACACACACACACCGCTAGCGATGCTCCATCTGTTGAGGTTTATGAAGACT 537
DB 37748 GCGAAGTCACTACATCTTTGACGAGATGAGTCACTTCCAGTCAATTTTCATGAAGTAT 37807
QY 538 TGGGCTTCCATCACTAAATTCGGAGAGATGATGAATTTCTTGCGGAAAGAGTGAATCT 597
DB 37808 TGGATGTCTATGACCAATTCAGCGGCAAGATCGCGCCACGCTTCTTCTACCTCTCTG 37867
QY 598 TTGCGCGTTTTCGACCGTCTGCTGTAATTTCCGCTTAATTTGACACATATTTATCG 657
DB 37868 CTTATTCAGCTGCGAAGAAACATAATCAAGGACCCAGGAGAGGTAGCCCGACGACATTTG 37927
QY 658 AACACGCGCAGAAACGTCGTTGGAATCGCAGCATCCATCTTTACCGACGATCGGATT 717
DB 37928 GAGCGGTTTGGAGCCAAACTCTGCAAAACACAGTTTCACATGTTACGCTGAGAACATG 37987
QY 718 C---GAGCTACTACCTTTTACCCCAACTGAAATTTAGAAATTTGAAGGTTTGTATTCAG 774
DB 37988 GTCAGAGCCACTTTTACATTGAGCGGGAAGCAGATAGACAATCTCAAGAGTTGGGTACA 38047
QY 775 AGAAAGCCCCAAATAGTAGTTAATCTCTTCTGCTGCGCATCGCACTTATATCTCG 834
DB 38048 GAGCAGCTTGAGAACTCAATCTCTGTTTCTACCTTCTGTTGCTGCTGCTTCAATTTG 38107
QY 835 ACCGCGATCGCAAAATCGGTCGGAGATTACA--AAGACGCTGGATGACCAACACGCGT 891
DB 38108 GTAAGCTTGATTAGACACTTTGTCAGACAGTGAACAGAGCGCAAGGCAAGTCTTT 38167
QY 892 TTCTTTTAAATTCGATGATTAAAGCGCGCTTTTGA-----TCCGCGGCTCGGGG 945
DB 38168 CACTTGATGATCAACGCTGATTGCGAGGATCTCTCAAGTACACACAACTACCAACAA 38227
QY 946 AACTACTTCGGAACCTGTATCTGTTTGGATGGGCAAGATCCTGCGGCGGATTTGGTC 1005
DB 38228 ACATACTTTGGCAACTGTATGGCTCTCTGTTATCGATCATGTCAGAAACACGATTTGCTA 38287
QY 1006 GGAGATGAAGGGGTGTTTTCGGGCACTGAGCGCATCGCGCGGCAATAGAGAGACG 1065

DB 38288 GGAGAAAAATCGTCTATGCTGCTCTCAGATGCAATCACAGCGAGAAATCAAGATATGTTA 38347
QY 1066 ACGCACAGAAGATCTCTAGAACTGTGGAGAACTGCGCCCTGTGAGATTCGCGAAGCCTTG 1125
DB 38348 TCAAGCGATCTCTTGAAGACAGACCAAGATGGGACAAAGGAGTACGTAATGGGTCTATG 38407
QY 1126 CAAAACCTGTTATTTCTCGTGGCGGATCGAGCAGCTTGTATCTTTACGCGCGGATTTT 1185
DB 38408 TCCATTATCCAACTCAATTTGAGGAGCTCCGAAATTTGGGACTGTATGACATGGATTTT 38467
QY 1186 GGATGGGTAAGCGGTGAAGCAAGAGATPACTGTTCGATTTGATGGAGAGAAGTTTACGATG 1245
DB 38468 GGATTAGGAAAGCCCTGCAAAATGGAGATAGTGCACATAGAAACAG--GTGGTTCTATC 38524
QY 1246 TCGTTGTGTAAACCGAGGATGCTCCCGGAGGATTCGAGGTTGGATTTCTTTGCCAAAG 1305
DB 38525 GCATTTCTAGATCTCCAGACGCGCAGCAATGAGTTGAGATTTGGAATAGCACTAGAGAAG 38584
QY 1306 GAGGAATTCGAAGCTTTTCATGATTTTTCGCGAGGGAATAAG 1350
DB 38585 AAGAAATGATGTGTTGACTCACTTCTACAAAGGGATCAAG 38629

RESULT 10
AC003027
LOCUS Arabidopsis thaliana chromosome I BAC F21M11 genomic sequence,
DEFINITION complete sequence.
ACCESSION AC003027
VERSION AC003027.1 GI:4079614
KEYWORDS HTG
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altai,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.S.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Kurtz,D.B., Conway,A.R.,
Au,M., Araujo,R., Buehler,E., Dewar,K., Feng,J., Kim,C., Li,Y.,
Oji,O., Osborne,B.I., Shinn,P., Sun,H., Toriumi,M., Vysotskaia,V.,
Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (22-OCT-1997) Biochemistry, Stanford University/DNA
Sequencing and Technology Center, 855 California Avenue, Palo Alto,
CA 94304, USA
REFERENCE 3 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altai,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (30-DEC-1998) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA
REFERENCE 4 (bases 1 to 119914)
AUTHORS Federspiel,N.A., Palm,C.J., Conway,A.B., Conn,L., Hansen,N.F.,
Altai,H., Araujo,R., Huizar,L., Rowley,D., Buehler,E., Dunn,P.,
Gonzalez,A., Kremenetskaia,I., Kim,C., Lenz,C., Li,J., Liu,S.,
Luros,S., Schwartz,J., Shinn,P., Toriumi,M., Vysotskaia,V.,
Walker,M., Yu,G., Ecker,J., Theologis,A. and Davis,R.W.
TITLE Direct Submission
JOURNAL Submitted (30-JAN-1999) DNA Sequencing and Technology Center,
Stanford University, 855 California Avenue, Palo Alto, CA 94304,
USA

COMMENT

On Dec 30, 1998 this sequence version replaced gi:2734094.
 Bases 1-9262 of clone F21M11 overlap with bases 68998-78259 of
 'TAMU' BAC clone F20D22 (AC002411) and bases 119525-119914 of clone
 F21M11 overlap with bases 1-389 of 'TAMU' BAC clone F21B7
 (AC002560).

e-mail for correspondence: arabidopsis.stanford.edu
 Genes with similarity to proteins in the databases are described as
 'putative', '-like' or 'similar to'. Genes that have EST
 similarity but no significant protein similarity are described as
 'unknown proteins'. Genes that are annotated based only on gene
 prediction software are described as 'hypothetical proteins'.
 The software programs used to predict genes include: Graal
 (informatics.group, Oak Ridge National Laboratory,
<http://compbio.ornl.gov/section/index.html>), GENSCAN (Chris Burge,
<http://compbio.stanford.edu/~chris/GENSCANW.html>), Fexa (V. Solovyev
 & A. Salamov, Sanger Centre, <http://genomic.sanger.ac.uk/>), and
 NetPlantGene (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, <http://www.cbs.dtu.dk/NetPlantGene.html>).

FEATURES

source
 1. 119914

/organism="Arabidopsis thaliana"
 /cultivar="Columbia"
 /db_xref="taxon:3702"
 /chromosome="1"
 /clone="F21M11"

misc_feature

1. 7000
 /note="Proteins in this region are annotated in the F21M11
 entry, AC002411."

misc_feature

1. 9262
 /note="Overlap with bases 68998-78259 of 'IGF' BAC clone
 F21M11."

gene

7684..10913

/gene="F21M11.1"
 join(7684..7866,7965..8027,8122..8160,8250..8381,
 8458..8728,8875..9893,10003..10256,10344..10410,
 10501..10551,10685..10781,10885..10913)

CDS

/gene="F21M11.1"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010665.1"

/db_xref="GI:4204284"

/translation="MAPNRIKKACDAMKLLGSETKTRAFRLKLLKTYENNNDFFIE
 DAYKVLDAIFADDAQSTENKKEEKKKRSVATSGRRKAPPELQVDEE
 DMDDEDFPLRLRRGRASSSSNNEDLKTQPEEDEDGVTLEPLKRY
 VRRNGERGLATVYNNAPSSSRLSMEPEEVPWLLPAHPWETKVSASALVIND
 EPNIDHPRVTSQDNCAPSAPMLEKSNHIVQEDWETKDLINDTANDVSPSSAIGES
 SEHKAASVELASSTSGEAKICLSFAPATGETTNLHPLSMEDLRAMEKCLRSYKI
 VHPFVSFLGPKMDCSCYIDLAKNSTQLLETETVCDMSKAGDESGAVGISPLVVP
 ECETSGDGRKAIKSNMKTITAGEENVEIPWNEINEKVPFRYMPHSFVQDAPVIFS
 LSSFDQSGSTCIEDCLASEMNCNCAIGVDNGFAVTLGLLKEEFLEARISEARDO
 RKQVLRFCESCPLEKAKKVEILEPCGHLKRGAKCEWKCCTKCGNRVQRGHN
 KLVQVFTPNKGWGLRTLEKLPKGAFCIEYIGEILTIPELYQSFEDKPLPVLDAH
 WGSERLEGKALCLDGMFYGNISRFLNHRCLDANLIEIPVQETPDQHYHLAFTT
 RDIEMAEALWDYIGIDFNNDLSMKPDCICGSRFCNRKRKSTKTMILNKA"
 join(13033..13488,13717..13926,14139..14514)
 /gene="F21M11.2"

mRNA

13033..14514

gene

/gene="F21M11.2"

join(13061..13488,13717..13926,14139..14316)

CDS

/gene="F21M11.2"

/note="Similar to acid phosphatase; Location of ESTs
 110C2T7, gb|T42036, and 110C2XP, gb|A1100245"
 /codon_start=1

/protein_id="A010666.1"

/db_xref="GI:4204285"

/translation="MDRTMFLSLTIASLLGVVSAGDWNILNQLGLGSSSSQNGIVS
 KGKTDLKGCESEINVEINRKKFDPVQECVSHIKDYMTSQYKDDVAVQDEVI
 LHFSGMCCSKCDGMDWIFDDITLLSTIPYHKNGFGGKLNSTFARDVTDQKK
 APAPVHKLYLHDIRERGIKFLISSKEVLRNTVDNLIOAGYGNLMLRGLEDO
 QKEVQYKSKRWLSLGRVWGMDQWSSGACPLPRATFLPNSIYYA"
 complement(join(14004..14312,14565..14641,14787..15831,
 15906..16006,16327..16362))
 /gene="F21M11.3"

mRNA

15906..16327

gene

complement(14004..16362)

/gene="F21M11.3"

CDS

complement(join(14567..14641,14787..15831,15906..16006,
 16327..16362))

/gene="F21M11.3"

/note="Unknown protein; Location of ESTs 40C3TT,
 gb|AA728590 and 40C3T7, gb|T04573"

/codon_start=1

/protein_id="A010667.1"

/db_xref="GI:4204286"

/translation="MLDQISGNHTHEKLSVETAEPHLNDNRVHIVEIPKASVPIPE
 ICDEAEKCSPTISRRRVTFDSKVITYEHVVSSELSSEKNEEVSRRKLSK
 TDQIIEVASNSTGSSYENHRYKNCRESDDIEDDCSDSLDEEYVSDYGFSE
 DSLHNPTEVYTDIGDKTEIDSKLRSTNETVRDGNHVGQGVLPNVENTOWKSAK
 SKGTQKQKQKENSFIADQEKRRDSFGTDIIDDITLSVKPKRICRPEKLNRQE
 LAVDASUTWLSTSESGECSASMYLTITPEKLSKSTSCYKPLRINHDDRPVLCALIL
 EDKQFSATSPTRKSPDPTIIGTVGGYWGNSKADICGSASSFKGIPNTSSKY
 REDKSVNHSHTPFARLEKALNNIDK"
 complement(16835..20238)
 /gene="F21M11.4"

gene

complement(16835..17185,17274..17392,17491..17656,
 17795..17885,17982..18079,18175..18361,18504..18604,
 18705..19049,19134..19349,19439..19791,19862..19967,
 20188..20238))

CDS

/gene="F21M11.4"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010669.1"

/db_xref="GI:4204288"

/translation="MGSSSPPEARAAQVPMILIFLEICTVHTVYNNRKLNRDVL
 NLNPKVTRDRKFRMESVTSYKSLAAVSVHLPQIPDNCYNKDALNNSPKH
 GESEDSMTDKVSKSGSDTSYSSRDCGSPLETSSEDPKHOPTKQLSDHLLYE
 FESEYDAANTHPTSEYTEQAANKNRDITASQPSNAARKRICGDSFIESSEPNKTD
 TLRLMESLRSDPTDIVKAOHOMVYFSFWIYFVLCRLNLRVCVIFVQOOLPK
 SHTEQSKRRDITASDAMENHLKVPKRENNLMQKSIDCNGKCSANSDDLSEKTS
 KALEQTSNITICGQSARVSEATGEMLHYSGRPVDDGDIIFRSNVIHVSACIEWA
 POVYEGDTVNKLKAEALGMKIKCTCSLKGALGCVKSCRSYHVPARETSRCR
 WDVEDFLLCPAHSSVFPNEKSHRYSRAEPLKINPAELCSLEQTPATKELVLCG
 SALSCKSLMESLAVRFNATISRYWNPSYTHVASTDEKACRTKLKVLGILNGK
 IIRNAHMKSLKASQPVDEEPEFIQIDTQCGQGPRTARLARAKTKPLFELGAFYF
 GEEYTKYKEDQLNLRKVAGGTILNDELGAESNNVNDQRSSIVVYINDDPPHGCAL
 GEYTIITWRANDAEALASQTSRLVGTWVLESYAGYKLPHPVIG"
 20752..24647
 /gene="F21M11.5"

gene

join(20752..20994,21362..21497,21596..21740,21825..21994,
 22102..22178,22542..22758,22920..23064,23343..23442,
 23599..23693,24021..24072,24227..24298,24441..24647)

CDS

/gene="F21M11.5"

/note="Hypothetical protein"

/codon_start=1

/protein_id="A010668.1"

/db_xref="GI:4204287"

/translation="MGANSKSVTASFTVIAVFFLICGRTAVEDETFHGDYKLSGI
 IIRFASTQLRAWSILDCPYTPLDNPDLVLDLDTTKLLSAVNCWFKCWLDPYNOTD
 HPCKSRPDSGLSAITELDPGYITPLSTVWKWKLVCEVEFIEANAIVAPYDMRLS
 PTKLEERDLYEHLKLTFTETALKLRGSPSVFAHSMGNVFRYFLEMLREIAKPHYL
 KWLDOHIAFYAVGAPLLGSVEAKLSLSTGVTGLPVSGETARLLSNSFASLLWMPF
 SKNKGNDTFTWHSFGGAARKDKRVHCEDEEYQSKYSWGMPNTIINIEISTARELA
 DGTILFKAIEDYDPSKRMHLQKKYVFFIRNIAHRSLSAGFLLYHDDPVFNPLTW
 ERPPKRVNFCIYGAHLKTEVGYFAPSGKPYPDNWIITDIIYETEGSLVSRSGTVVDG
 NAGPITGDETVPYVSHLSWCKNLGCPKNITMAPQILGKIKQOPEHSGSDVHVELNVD
 HEHGSIIANNKAPRVKITYTFEEDSESIPGKRTAVWELDKGY"
 27777..28734
 /gene="F21M11.6"

mRNA

27777..28734

gene

27777..28734

CDS

/gene="F21M11.6"

28007..28465

/gene="F21M11.6"

/note="Unknown protein; Location of ESTs 203I24T7,
 gb|H76794 and 203I24XP, gb|AA605510"
 /codon_start=1
 /protein_id="A010670.1"
 /db_xref="GI:4204289"

/translation="MNTKMRLLPPRRVLITADKKRERDAFISSVTDNPPPIAKFPSPPP
KLPPVPNPISKSSATAAEPIGSNQLMLAGYLSHEYLTQGTFLFEQNNQARAQESS
KIKPSHTVPEAECEPKRKRYREVANLLRSQAQLPGIVNPAQLARFLK"
complement(29264..32033)
/gene="F2LW1.7"
complement(join(29264..31015,31312..31414,31484..31587,
31782..32033))

Query Match 5.8%; Score 87.8; DB 12; Length 119914;
Best Local Similarity 44.9%; Pred. No. 7e-08;
Matches 617; Conservative 0; Mismatches 717; Indels 39; Gaps 6;
QY 11 TCCTCGAATCTCCCGAGTGGCGCCCTCCAGGACAGGTGGCTGAGCAGTCACTCCCGC 70
DB 45957 TCATCGAGACCTGTACATCTCTCCCAAGGCGACCGTCCATCAACCACCTCTCCCTC 46016
QY 71 TCACCTTCTTCGACATGAGCTGGCTGCATTTCCACCCCATGCTTCAGCTTCTCTCTAGC 130
DB 46017 TCACCTTCTTCGATGCCCTCGCTCTCTCCCACTCCCGCATCTCTCTCTCTCTCT 46076
QY 131 AACTCCCTGTTCACAAACCGCTTCTTCGAAACCGTCTTCGGAACCTCAACAACTCCT 190
DB 46077 CTTACCAAACTCACTGAATCTTCTCTCCAGACHTTGTACCAACCTCAACAACTCCC 46136
QY 191 TATCTCTAACCTCAACACATCTTCTCCCTCTTTCATGCAATCTAATCTACCTCTATCGC 250
DB 46137 TCTCCATCACTCTCCAACATTTCTTCCCTTACGCCGTAAACTGATTATCCCGCCTCGT 46196
QY 251 CGGAGAAATGCGGAGTTCCGGTATCAGAACGGTGACTCGGTTCTTTCACGATTTATGG 310
DB 46197 CGGACCTCCATATTTACACTACACGATGGCCAAAGACTCTCTTGTTTTACCGGTAGCAG 46256
QY 311 AGTCT-----GTCGGAGATCATCCGATTCGCTCATAAATACTACTGCTTTGGCC 361
DB 46257 AGTCTACTGAACCGACTTCGACCACTGAAGTCTGATTCCTCCCTAAGATATCAGTGTG 46316
QY 362 CTAGCGACGATTAAGAATCTCCAGCTCCGCCGAGTCTGAGGAAATCTGATCGGAAAT 421
DB 46317 TGCATGGCTGTGGCCAAAGTTACCTCTCTCAGCTCTCTCCGGAAGAAATTCAGA--- 46373
QY 422 TGTTCGAAGTTTAGCCGTGCAAGTACTCTGTTTCCCGTCCGGGTGTCATCGGA 481
DB 46374 TCGGACCAATAATGGCTATGCAAGTACCACTCTTCCCTGGAGCTGGAATCTGTATAGGA 46433
QY 482 TAACGACGACACACCGTTAGGAGTCTCCATCTGTTAGGTTTATGAAGAGTTGG 541
DB 46434 ACTCAGCTACACATGTTGTAGCAGATGGAGTCACTTCACTATTCATGAAGTATGGA 46493
QY 542 CTTCCATCACTAAATTCGGAGGAGATGATGAATCTTTGGACGGAAGGTGAATGTTGC 601
DB 46494 TGTCAATTGACCAATCCAGCGGTAAGATCCGCCACCGTCTTCTACCTCTCTCGCCTA 46553
QY 602 CGGTTTTCACCGATCGCTCGTAATTTATCCGGCTAAATTTGGACACATATTTATGGAACA 661
DB 46554 TTCACAGCTGCAAGAACATGATCAAGGACCCAGCGGAGGTAGCGGACGACATTTTGAGC 46613
QY 662 ACGCGCAGAAACGTCGTTGGAAATCCAGATCCATCTTTTACCGCGGATCGGATTC--- 718
DB 46614 GGTTTTGGAGCCAAAACCTCTGCAACACAGATTCACATGTTAGCCCTGAGAACATGGTCA 46673
QY 719 GAGCTACCTACCTTTTTCACCAATCTGAATTAAGAAATTTGAAGGTTTGTATTCAGAGAA 778
DB 46674 GAGCCACTTTTACATTTGAGCGGAAGCAGATAGATAATCTAATAAGTTTGGTTTACAGAGC 46733
QY 779 AAGCCCAATATGATTAATCTCTCTCTCTCGTCCGGATCGCAGCTTATATCTGACCG 838
DB 46734 AGTCTGAGAAATCAATCTCTGTTTCTACCTCTGCTGGTGACTCTAGCTTTTATTTGGGTAA 46793
QY 839 GCATCCCAATCGGTGGAGATTA-----CAAGACGTGGGTGACGACA 883
DB 46794 GCTTGATCAAGACACTTGTACAGACAGTGAACAAAGGCCAACGAGGAAGACAAGATG 46853
QY 884 AACGCGCTTTCTTTTAAATTCGATCGATTTAAGGCGCGGTTTGGAT-----CCGCGG 937

DB 46854 AAGTATTTTCACTTGATGATCAATGTGGATTGAGGAATCGCTCAAGTACACACACCAA 46913
QY 938 CTCGCGGGAACACTTCTCGAAACTGTCTATCGTTTCGATGGGGAAGATCCTCTCGCGCGG 997
DB 46914 TACCACAAACATACTTTGCAACTGATGCTCTGCTGCTGATCAGTCAAGAAACACG 46973
QY 998 ATTGTGTCGAGATGAAGGGTGTTCGCGGACGCTGAGCGATCGCGGGGGAATAGAGA 1057
DB 46974 ATTGTCTAGGAGAAATCGTTTTGCGCGCTTCAGATCAATCAACGAGAGATCAAAG 47033
QY 1058 AGAGGACGAGCAGAGAAGATTTAGAAACTGTGGAGAACTGGCCGCTCTGAGATTCGCG 1117
DB 47034 ATATGTTATCAAGCGATCTGTTGAAGACAGCACCAAGATGGGACAAAGTAGCTAAAT 47093
QY 1118 AAGCCTTCGAAACCTGTTATTTCTCGTGGCGGATCGAGCAGCTTGTATCTTTACGGG 1177
DB 47094 GGTCTATGTCCCATATTCACCTCAATTCGAGGAGCTCCGAAATTTGGAGCTGTATGACA 47153
QY 1178 CGGATTTTGGATGGGTAAAGGGGTGAAGCAGAGATCTGTCTGATTTGATGAGAGAAGT 1237
DB 47154 TGGATTTTGGATTAGGAAGCCCTGCAAAATGGAGATCGTCACATAGAAACAG---GTG 47210
QY 1238 TTACGATGCTGTTGTATAACGAGGAGTCTGTCGCGAGAGTTGGAGTTGGATTTGCTT 1297
DB 47211 GTTCTATCGATTTCTCAGAGTCCAGAGACGGCAATGGAGTTGAGATTGGAATAGCAC 47270
QY 1298 TGCCAAAGAGGAAATTCGCAAGCTTTTGTGATGATTTATTTTCGGGAGGGAATAAG 1350
DB 47271 TAGAGAAAGAATAATGGATGTGTTTGACTCAATTTACAAACAGGGGATCAAG 47323
RESULT 11
AB028618
LOCUS AB028618 85690 bp DNA PLN 27-DEC-2000
DEFINITION Arabidopsis thaliana genomic DNA, chromosome 3, Pl clone:MOD1.
ACCESSION AB028618 BA000014
VERSION AB028618.1 GI:5041971
KEYWORDS Arabidopsis thaliana (strain:Columbia) DNA, clone_lib:Mitsui P1
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
Brassicales; Brassicaceae; Arabidopsids.
REFERENCE 1 (sites)
AUTHORS Kaneko,T., Katoh,T., Sato,S., Nakamura,A., Asamizu,E. and Tabata,S.
TITLE Structural analysis of Arabidopsis thaliana chromosome 3. II.
JOURNAL Sequence features of the 4,251,695 bp regions covered by 90 Pl, TAC
MEDLINE and BAC clones
REFERENCE DNA Res. 7 (3), 217-221 (2000)
20363099
REFERENCE 2 (bases 1 to 85690)
AUTHORS Kaneko,T., Kato,T., Sato,S., Nakamura,Y., Asamizu,E. and Tabata,S.
TITLE Direct Submission
JOURNAL Submitted (09-JUN-1999) Yasukazu Nakamura, Kazusa DNA Research
Institute, Department of Plant Gene Research; 1532-3, Yana,
Kisarazu, Chiba 292-0812, Japan (E-mail:ynakamu@kazusa.or.jp,
Tel:81-438-52-3935, Fax:81-438-52-3934)
COMMENT Address for correspondence: kaos@kazusa.or.jp
For the latest information on annotation of this clone, please see
http://www.kazusa.or.jp/kaos/cgi-bin/agd_graph.cgi?c=MOD1
Genes with similarity to proteins in the databases are described in
'product' or 'note' qualifiers. Genes that have no significant
protein similarity are described as 'unknown protein'.
The software programs used to predict genes include: Grail
(Informatics Group, Oak Ridge National Laboratory,
http://compbio.ornl.gov/Grail-1.3/),
GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of
Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and
SplicePredictor (Volker Brendel, Stanford University,
http://gremlin1.zool.iastate.edu/cgi-bin/sp.cgi).

Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is T1310 and the 3' clone is T26612.

FEATURES

source

```

1. 85690
/organism="Arabidopsis thaliana"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="3"
/clone="MOD1"
/clone_lib="Mitsui p1"
2123. 8637
/note="contains similarity to TNP2-like transposon
protein
gene_id:MOD1.1"
/codon_start=1
/pseudo
/evidence-not_experimental
join(10671..10844,10921..11208)
/note="contains similarity to unknown protein
emb|CAA16722.1
gene_id:MOD1.2"
/codon_start=1
/pseudo
/evidence-not_experimental
15142..16497
/note="gene_id:MOD1.3"
/codon_start=1
/evidence-not_experimental
/product="anthocyanin 5-aromatic
acyltransferase/benzoyltransferase-like protein"
/protein_id="BAB02518.1"
/db_xref="GI:11994477"
/translation="MTLHVETARVTPTYSVINSANLKLPLTFDPLWLLFQPVKR
VFFVELTSDRDHFHSIILPKLDSLILRNLYPLTGHITWPNPKPSIIYSENGV
VLVTIAESDAFSLGSGYRGORPLSELHALVPLKPVSDSATAFSIOITLFPNGFSIG
VAHHAVLDCKTSFTFKAWAQICKQELQSPENLTPSYDRSLIKPTYIDKRIELY
RSLKEDVTNLRSLTSLPSKLGDDVLATLVLSRADIERUEQVKNVSPSLHLSTFVI
AYAWTFVKRGKNDKRSVLLVFGDFRDLDPKLPGTGYFCNCPMPVGCYNRKAEE
FMEKGVTAAEIIISLVKGLSKRIETIADTFVEGFSQSWSTQFGTIGASTRLGVY
EADFGWRPVKVDIVSDQGEATAMARRDESQVEIGMCKLTKTMDSVSVFFNGLH
S"
complement(18639..19994)
/note="gene_id:MOD1.4"
/codon_start=1
/evidence-not_experimental
/product="anthocyanin 5-aromatic
acyltransferase/benzoyltransferase-like protein"
/protein_id="BAB02519.1"
/db_xref="GI:11994478"
/translation="MALNVIKISRVSLVNSVEPLVPLPTFFDOLLMLKNPIERVTFY
KLTESSRDSFSLKLEQSLVLSHLFPLSGHLKNPQDPKPHIVIPKPTVSIT
VWSEADFSYISSKELRELTELRLPELVSDSASLLSLOITLFPNGFSIGTIVH
HYVMGDKASKFHKWAHICKHGTPODEDLPTVLDRTVINVPAGLEOKIFOLSSYIS
EKDYATPLLPKAKELDNDVVRVLTLETVDIEKLERAKNESTRSDLHLSTFVSY
AYVLTKVWCKCGGANRPVRPMAADFRNRDPPVLTLYFGNCVLPIDFNGYKATITFL
GKGYGVNVEILSDVSGLSRNIESIWEYEDGTKNMKLDQNTVITVTSNQFGIYGS
DFGWRGVKVDVMSLYKNNEFSMSARDEIGGLEIGTISLKKCNMNVFLSLTSDFDIY
K"
complement(27960..28438,28676..28739))
/note="gene_id:MOD1.5"
/unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB02520.1"
/db_xref="GI:11994479"
/translation="MRLAGACRSPSEESRRRRRARAGWSVDRPIRFIYAADFRNRD
PPVLTITFGNCVLPIDFNGYKATITFLGEGDFVNGVEILSDVSGLSRSLSESVWEI
EGTKNMKGVTGVITVTSNQFGIYAGDFGWRGVKVDVMSLYKNNEFSMSARDEITGV
SLKRCRMIVFLSLFNGFDN"

```

CDS

```

29217..31674
/note="gene_id:MOD1.6"
/codon_start=1
/pseudo

```

```

/evidence-not_experimental
/product="Mutator-like transposase"
join(32547..32726,32904..33122)
/note="gene_id:MOD1.7"

```

CDS

```

/codon_start=1
/evidence-not_experimental
/product="Mutator-like transposase-like protein"
/protein_id="BAB02521.1"
/db_xref="GI:11994480"

```

```

/translation="MDRLCRDPYDDMKVAKRAIEOMVAMMEGIPKFCPCGGSTV
DTRKDEKYYQCEKEDDTCDHKIRKLWDKAMEEEVSSLRESVDYNRKVKLSHEYLI
EEMQKELAHRAEIVNVSKVLFRPNHAPKK"
complement(join(33460..33573,33657..34016,34094..34255,
34330..35070,35486..35721,35799..36093,36146..36256))
/note="gb|AAC23763.1
gene_id:MOD1.8

```

CDS

```

strong similarity to unknown protein"

```

```

/codon_start=1
/evidence-not_experimental
/protein_id="BAB02522.1"
/db_xref="GI:11994481"

```

```

/translation="MPPKTRGGQGRKNEIEASAPAKTEKVKAPAEKVKKEGAPSKAK
VOAPAKKAVQAPAKKDEKVPAAEQSPQTTATAMATNAAPTAAPTTAPTAPTPT
ESPLDDSTFYDALKHTPAEIOENMQTDEVEDEETENKELACANVEEAERQDDGL
TVIEEERSESEDEDVNNKSVSEDEGHEDERDVIKESGEERTIDEDIANVDMEE
AMAMQPLMGTFYPSVEYPPKMKLATRCYISEVLKTFADLEHLTHVEKNYFMEHPSFKH
IYHLPQSYTHKLMRMMLFRTKGTIRREDVKTLMEMEPARSKDRLRMAALYFLTIVMP
PASYQSAGNNFANFYKGTGIRREDVKTLMEMEPARSKDRLRMAALYFLTIVMP
TKTGERASPIDDCVRSASDLTFCKIIPPGRGYSEYMLKSIHKLDFHNGVVPNTQSP
WVPVGCVPLEFLAFAEAPLSRERFIEEKAGACGCMCKVNFKRLKMGFFLEQIN
HVLGTVEISIIIRAEAPVLLAEITGVEDDKHDVVVDVMEKLHSINDKLEALLVLMIEBKQAT
FEAFDMQMAKMSQNPDEESTIKENAAAPVVPKRVTRSTRASNSV"
complement(36608..37000)
/note="gene_id:MOD1.9"
/unknown protein"
/codon_start=1
/evidence-not_experimental
/protein_id="BAB02523.1"
/db_xref="GI:11994482"

```

CDS

```

/translation="MCKTASKEVKAHICKHGNIPQDFDLPMLDRTVINVPAGLE
PKLFQLLPYLSKEKVARNMLPPAKENINVRVTLESEANIKKLKEQAKNESTWSD
LLISTFVVTYKSSGSLFCKEEAKKSL"

```

```

complement(join(40204..40603,40656..40872))
/note="gb|AAC97247.1
gene_id:MOD1.10
similar to unknown protein"
/codon_start=1
/pseudo

```

CDS

```

/evidence-not_experimental
complement(join(42135..42706,42772..43162))
/note="gene_id:MOD1.11"
/unknown protein"

```

```

/codon_start=1
/evidence-not_experimental
/protein_id="BAB02524.1"
/db_xref="GI:11994483"

```

CDS

```

/translation="MALNMIKISRVSHATNSVTPLELPTFFDOLLMLKHHPIERVTFY
KLSESSHDSFFSTILPKLEQSLALTNFLPSLGHIKWDPQPKPHIVIFPDQTVSS
LWLSPLTSDSPAKSLVSRSYVLLSLRSLGTHHHVIMDKTSTNFHKSNAHICK
LVRTIPLQGVLPVTHSDRTVINVSGLFEKILQLLSYLSGVDNDRTILKSPSAKEIDD
LVRTIPLQGVLPVTHSDRTVINVSGLFEKILQLLSYLSGVDNDRTILKSPSAKEIDD
FYAAADFNRNLDPPVLTLYFGNCVLPIDIFHGRRWCEWRGDSK"
44134..44573

```

CDS

```

/note="gene_id:MOD1.12"
/unknown protein"
/codon_start=1
/pseudo
/evidence-not_experimental

```

[illegible]

```
Db 91369 TCTTTTCAACAATGGTTTACATAGCACTCTGTTTGTGAAGATTTTAATAAAG 91428
Qy 1387 GATGAATCCTCTGTTTCATCTCTATGTTTAAACAATAATTTTTCATTTGAACCTTTT 1446
Db 91429 GTAATGCTTGGTTAGTATTCTGTTTCTTCTCACAAGAGTATTGGCATAATCTCC 91488
Qy 1447 TTGAGTCAATAAAAAAATAAAAAA 1475
Db 91489 ACTAATAAATAGTTTAAATAAGATATA 91517

RESULT 13
AX025514
LOCUS AX025514 1648 bp DNA PAT 16-SEP-2000
DEFINITION Sequence 40 from Patent WO0032789.
ACCESSION AX025514
VERSION AX025514.1 GI:10187184
KEYWORDS Citrus limon.
SOURCE Citrus limon.
ORGANISM Citrus limon.
REFERENCE 1 (bases 1 to 1648)
AUTHORS Aharoni,A., Verhoeven,H.A., Luecker,J., O'Connell,A.P. and Van Tunen,A.J.
TITLE Fruit flavour related genes and use thereof
JOURNAL Patent: WO 0032789-A 40 08-JUN-2000.
AHARONI ASAPH (IL) ; VERHOEVEN HARRIE ADRIANUS (NL) ; LUECKER JOOST (NL) ; CPRO DLO (NL) ; CONNELL ANN PATRICIA O (NL) ; TUNEN ARJEN JOHANNES VAN (NL)
FEATURES
source Location/Qualifiers
1. .1648
/organism="Citrus limon"
/db_xref="taxon:2708"
52. .1527
/note="cDNA"
Lemon acyl transferase"
/codon_start=1
/protein_id="CAC09066.1"
/db_xref="GI:10187185"
/translation="MAAIENKRVLUKHEVTKVPFVNPNSKTSFTLIDLTYFFFWPK
NPVERLFFVDFWDLNSENLEPKLHSLFTLHLPLAGHIMPLDAAKPAVY
YFPDQGVFAVAENWSECHGPHLSGIRQAVEFHLVPLQSLSDKAEVIAIQ
ITLPNGGFSIGVSHHAILDQSTFLKSWAYLCKLOLCHHPCLSPELTLLDRT
VIKPTQDMLQNLNVGVSDNSDPQKIRSLKVLPLFJDSLNLKHPCLSPELTLLDRT
KLRLKVNHLKSKSKQVRLSTFLVFLAYVFCMAKAKAKATEAAAGANDEIKN
LIIVGFTADYRSRDLPLPLNIFGNCNGRHCTAKASDFVQENGVAFAEMLSDMKVI
DADAIEDDKVSEILKLGAMIFSVAGSTQPDVYGSDFGWRGPRKKVEIVSIDRTQ
AISAERDGGGVEGVVLEKQOMEFESVFADGLKNDLV"
BASE COUNT 464 a 344 c 354 g 486 t
ORIGIN

Query Match 4.7%; Score 71.2; DB 9; Length 1648;
Best Local Similarity 56.2%; Pred. No. 0.00012;
Matches 199; Conservative 0; Mismatches 143; Indels 12; Gaps 3;

Qy 1137 TTTCTCGTGGCGGATCGACGAGCTTGATCTTTACGGCGCGATTTTGGATGGGTAA 1196
Db 1302 TTTTCTGTGCTGCTGCACCCCAATTTGATGTTTACGGTTCGGATTTTCGGGTGGGGAG 1361
Qy 1197 GCGGTGAAGCAAGATACGTCTGATTGATGGAGAGATTTTACGATGCTGTTGTGTA 1256
Db 1362 GCCCAAGAGGTGAGATTGTCAATAGATAG---GACACAAGCCATCTCTTTGGCAGA 1418
Qy 1257 ACCGAGGATGCTCGCGGAGGATGGAGTTGGATTTGCTTTGCCAAGAGGAGTGA 1316
Db 1419 GAGAAGAGATGAGGAGCGGCTTGAGGTTGGAGTTCTTTTAGAAGAACCAAAATGA 1478
Qy 1317 AGCTTTTGATGATTATTTTGGAGGGAAT-ANAGGGTTGATTAATCATTTAATCATGA 1375
Db 1479 GGTTTTGAATCTGATTGCTGATGGGACTGAAAAAATGATCTTGTGTTAATGATGTA 1538
```

```
Qy 1376 TTATGAAGTTGGATGAATCCTCTGTTTCATCTCTATGTTTAAACAATAATTTTTC 1435
Db 1539 TCAT-----CTAAATTTCTCAATATATTATTTGGTCATATTCAAAGAAATAAATTAT 1590
Qy 1436 ATTGAACHTTTTGTAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1489
Db 1591 TCGGATTTTGTGACCAACCAATAAAATACTCTTTTGTGAAAAAATAAAAAA 1644
```

```
RESULT 14
CNS07360
LOCUS CNS07360 1167 bp DNA STS
DEFINITION clone BAOAB017A09 of library BAOAB from strain CLIB 210 of Kluyveromyces lactis, sequence tagged site.
ACCESSION AL427102
VERSION AL427102.1 GI:12210296
KEYWORDS STS.
SOURCE Kluyveromyces lactis.
ORGANISM Kluyveromyces lactis.
REFERENCE 1 (bases 1 to 1167)
AUTHORS Bolotin-Fukuhara,M., Toffano-Nioche,C., Artiguenave,F., Duchateau-Nguyen,G., Lemaire,M., Marmeisse,R., Montrocher,R., Robert,C., Termier,M., Wincker,P. and Wesolowski-Louvel,M.
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 11.
JOURNAL Kluyveromyces lactis
PUBMED FEMS Lett. 487 (1), 66-70 (2000)
REFERENCE 11152886
AUTHORS Souciet,J.L., Aigle,M., Artiguenave,F., Blandin,G., Bolotin-Fukuhara,M., Bon,E., Brottier,P., Casaregola,S., de-Montigny,J., Dujon,B., Durrens,P., Lepingle,A., Llorente,B., Maupertuy,A., Neuveglise,C., Ozier-Kalogeropoulos,O., Potier,S., Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M., Wincker,P. and Weissenbach,J.
TITLE Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of yeast species for molecular evolution studies(1)
JOURNAL FEMS Lett. 487 (1), 3-12 (2000)
PUBMED 11152876
REFERENCE 3 (bases 1 to 1167)
AUTHORS Genoscope.
DIRECT SUBMISSION
TITLE Submitted (08-Sep-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)
JOURNAL This STS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.
```

```
FEATURES
source Location/Qualifiers
1. .1167
/organism="Kluyveromyces lactis"
/strain="CLIB 210"
/variety="lactis"
/db_xref="taxon:28985"
/clone="BAOAB017A09"
/clone_lib="BAOAB"
1. .1167
/note="part of mitochondrial DNA"
/evidence=not_experimental
BASE COUNT 335 a 117 c 96 g 498 t 121 others
ORIGIN
```

Query Match 4.6%; Score 70.2; DB 53; Length 1167;

KRCIGQKHPPAPYTRDAQGLWDHLTTSRRHHKTLFKVSHQHSKLLVCKGVGRVLEYSNS
 ELMRLMPGPTVWNSVEDGSDRSKHKDKCIRTHPIRAHGCNLGTGKGR
 KHQWQVNRTHHGISIRLPSYCSLLSHGETDRRSVRPNGSDMDNQRESRPWPMQKT
 DVGHASKPKADGRISWASLR"
 join(25891, .26085, 26155, .26349, 26412, .26534)
 /note="contains similarity to Mutator-like transposase
 gene_id:MT024.8"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="BAB01194.1"
 /db_xref="GI:11994165"
 /translation="MDRLCSDDPHYDEMKSARKIMQOLEEVANMEETPIIPCQGGRI
 DIISEKQDGKRIYECTYDNGDLGHQKLWDKAMVEENRKLREQVDNHHOKTOSLEY
 SNQVLSSEFDIQKMKTLWRVRYWCYYRYQVHCILSIPLYSQFSLVFVLSLNL
 SFPLYSLYR"
 complement(join(26878, .27120, 27210, .27491, 27693, .27854,
 27950, .28175, 28419, .28702, 29026, .29487))
 /note="emb|CAB57796.1
 gene_id:MT024.9
 strong similarity to unknown protein"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="BAB01195.1"
 /db_xref="GI:11994166"
 /translation="MARPTIPLKRRKGVAGTKPIDVVTKATPTPTTTEPSATE
 QNPVAEGGVEEPIPTIVVEGQSDNNESEKDENEESESESESESESESESESE
 EKEEKEENVEGNTTGGSSDSTTLEEYSDDKNMNDATAVENQVEIPASMEIDEET
 KAILPLGPPSPVVKIKLSIRCYTHEVLTTFDKLEPMSKSEREFQNHPSQDHT
 FHMPDRNHLRMLGMMLLTARTERKEACIIAPLKTGDKAPOVDFCLKAVSOLT
 RETFMQGRFFDYMLGTISHTVNHNGSVTNEKYSVSVGLFCLOMELLAFAFAIPL
 RETFMQIADAGCCPRMCVRKKNHLKGFPDLTIYAELGICQRRKSELAFAIFEI
 QQDIALQITNAPKIAAGDIPAVGGQELFSVSVEMENLDRKMVDQLEIITMIS
 QDKRVSELAQDEQKAEKRNQOEMDQKDGDPIDKGVQKDGQEDGVPKEKQ
 QEDKGVKPKDPEKDEKDGKSPKVPVEANFTRITRTTKN"
 join(33399, .33608, 33762, .33830, 34025, .34098)
 /note="gene_id:MT024.10
 unknown protein"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 join(34168, .34242, 34530, .35934, 36128, .39479)
 /note="gene_id:MT024.11"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 /product="retroelement pol polyprotein-like"
 39663, .40357
 /note="gene_id:MT024.12
 unknown protein"
 /codon_start=1
 /pseudo
 /evidence=not_experimental
 join(49318, .49683, 49880, .49959, 50068, .50400, 50660, .50679,
 50779, .50980)
 /note="contains similarity to Mutator-like transposase
 gene_id:MT024.13"
 /codon_start=1
 /evidence=not_experimental
 /protein_id="BAB01196.1"
 /db_xref="GI:11994167"
 /translation="MKRNVDSTFKLRVSPFYSFOGKRKLRPNYIRDNEDLCYFFEDV
 NEEGFRSILHVESNWRNENRDPKILREDLRQNVNNEEDTSIIVGTDEIGAVLAM
 YVGSEVVENKDIYLEINICGMSKPYSIKRDSPMSVKNYRAKFRVQVSYSTAM
 KGRKAAANDKVTPEEGFRVDYMYLMQNHESGSYVEVDENFNLYLFFALGSSI
 EGFRLMKVILIDGTHKFRYAGDIIIVATQDPVHRCASPCWECARAYTEAEFLRGY
 GGEKYNROTNSVESVNGVLEKVRNYSLLQIDAIVGKIAEWFAKHKSLLELPSGY
 LDPRD"
 join(58702, .58758, 59228, .59328, 59533, .59839)
 /note="gene_id:MT024.14
 unknown protein"
 /codon_start=1
 /evidence=not_experimental

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:11:58 ; Search time 445.49 Seconds
(without alignments)
2145.203 Million cell updates/sec

Title: US-08-894-356C-6

Perfect score: 1522

Sequence: 1 NTGACCACCCCTCGAATC.....TTTTTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 10%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_0601.*

1: /SIDS1/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDS1/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDS1/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDS1/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDS1/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDS1/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDS1/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDS1/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDS1/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDS1/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDS1/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDS1/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDS1/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDS1/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDS1/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDS1/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDS1/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDS1/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDS1/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDS1/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDS1/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDS1/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1505	98.9	1518	17	AAT37313 Aromatic acyl tran
2	765.4	50.3	1479	17	AAT37311 Aromatic acyl tran
3	196.2	12.9	1703	17	AAT37308 Aromatic acyl tran
4	174.2	11.4	1508	17	AAT37312 Aromatic acyl tran
5	153.8	10.1	1622	17	AAT37309 Aromatic acyl tran
6	87	5.7	936	22	AAF58252 Oligonucleotide D1
7	87	5.7	936	22	AAF58254 Oligonucleotide D1
8	87	5.7	936	22	AAF58257 Oligonucleotide D1
9	87	5.7	936	22	AAF58259 Oligonucleotide D1
10	87	5.7	936	22	AAF58262 Oligonucleotide D2
11	87	5.7	938	22	AAF58255 Oligonucleotide D1

12	80.2	5.3	936	22	AAF58252 Oligonucleotide D1
13	80.2	5.3	936	22	AAF58254 Oligonucleotide D1
14	80.2	5.3	936	22	AAF58257 Oligonucleotide D1
15	80.2	5.3	936	22	AAF58259 Oligonucleotide D2
16	80.2	5.3	936	22	AAF58262 Oligonucleotide D2
17	80.2	5.3	938	22	AAF58255 Oligonucleotide D1
18	72.4	4.8	244	22	AAF58238 Oligonucleotide D1
19	72.4	4.8	244	22	AAF58238 Oligonucleotide D1
20	71.2	4.7	1648	21	AAC64785 Lemon acyl transe
21	65	4.3	1240	21	AAC53227 Arabidopsis thalia
22	62.4	4.1	49999	20	AAC23902 Human LOBO homolo
23	61.6	4.0	1379	21	AAC46852 Arabidopsis thalia
24	55.4	3.6	49999	20	AAC23902 Human LOBO homolo
25	55	3.6	2418	13	AAQ27886 P.falciplarum GBp13
26	54.6	3.6	3030	21	AAC61886 cDNA encoding a hu
27	54.2	3.6	1520	21	AAF15829 Human prostate can
28	54	3.5	2652	10	AA90703 Rhoptry membrane a
29	52.6	3.5	1875	18	AAT48669 Human EDG-2 recept
30	51.8	3.4	1213	21	AAC51043 Arabidopsis thalia
31	51.8	3.4	1217	21	AAC42550 Arabidopsis thalia
32	51.4	3.4	3138	12	AAQ11712 Shuttle vector pMU
33	51	3.4	8365	20	AAZ20056 Plasmodium falcipa
34	50.8	3.3	1052	10	AA90224 Malaria-specific D
35	50	3.3	1518	17	AAT37313 Aromatic acyl tran
36	50	3.3	5173	18	AAT89783 Human Immunodefici
37	49.8	3.3	1132	20	AA80668 Clone am996_12 enc
38	49.8	3.3	2229	21	AA23425 cDNA encoding huma
39	49.6	3.3	3138	12	AAQ11712 Shuttle vector pMU
40	49	3.2	2564	22	AA500034 Human ATLAS-3-enco
41	49	3.2	4590	7	AA60472 Sequence encoding
42	48.8	3.2	2418	13	AAQ27886 P.falciplarum GBp13
43	48.6	3.2	2535	21	AAF18128 Lung cancer associ
44	48.4	3.2	1493	11	AAQ04492 Sequence encoding
45	48.4	3.2	1493	20	AA33994 Human osteopontin

ALIGNMENTS

RESULT 1
AAT37313
ID AAT37313 standard; cDNA to mRNA; 1518 BP.
XX
AC AAT37313;
XX
DT 06-FEB-1997 (first entry)
XX
DE Aromatic acyl transferase coding sequence.
XX
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KW plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla oclmoides;
KW Scenecio cruentus; Lavandula angustifolia; ds.
XX
OS Lavandula angustifolia (Clone pLAT21).
XX
FH Key Location/Qualifiers
FT CDS 1..1355
FT /*tag- a
FT /product= Aromatic acyl transferase.
XX
XX WO9625500-A1.
XX
XX PD 22-AUG-1996.
XX
XX PF 16-FEB-1996; 96WO-JP00348.
XX
XX PR 30-JAN-1996; 96JP-0046534.
XX
XX PR 17-FEB-1995; 95JP-0067159.
XX
XX PR 29-JUN-1995; 95JP-0196915.
XX
XX (SUNR) SUNTORY LTD.

PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
PI Nakao M, Tanaka Y, Yonekura K;
XX WPI; 1996-393401/39.
XX
XX DNA coding for aromatic acyl transferase - for transforming plants
PT which produce anthocyanin pigments and thus altering colour tone,
PT e.g. of flowers
XX
XX Claim 4; Page 73-76; 94pp; Japanese.
XX
XX Vectors containing DNA fragments encoding proteins of plant origin
CC with aromatic acyl transferase activity may be used to transform
CC plants which produce anthocyanin pigments. The aromatic acyl
CC transferase acylates the pigments in the flower resulting in colour
CC tone changes and allowing new colourations to be produced. Six
CC specific DNA sequences encoding aromatic acyl transferase from
CC different plants are described in AA137308-137313. NOTE: This
CC sequence is supposed to cross reference with the protein described
CC in AA04727, however there are so many discrepancies between the
CC polypeptide decoded from this sequence and the polypeptide given in
CC the specification and described in AA04727 that the indexer decided
CC not to cross reference the two.
XX
XX Sequence 1518 BP; 384 A; 340 C; 366 G; 428 T; 0 other;
SQ

Query Match 98.9%; Score 1505; DB 17; Length 1518;
Best Local Similarity 99.8%; Pred. No. 1.2e-278;
Matches 1518; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 2 TGACCACCCCTCCTCGAATCTCCCGAGTGGCGCCGCCCTCCAGGACAGGTGGCTGAGCAGT 61
DB 1 Tgaccacccctcctgaa---tcccgagtggcgccgccctccaggcaagtggtgagcagt 57

QY 62 CACTCCCGCTCACCTTCTTCACATGACGTGGCTGCATTTCCACCCCATGCTTCAGCTTC 121
DB 58 cactcccgctcaccttcttcacatgacgtggctgcattccaccccatgcttcagcttc 117

QY 122 TCTTCTAGCAACTCCCTGTTCGAACCCGCTTCTCGAAGCGTCTTCGGAACATCA 181
DB 118 tcttctagcaactccctgttccgaacccgcttcttcgaacccgcttccgaacatca 177

QY 182 AACAACTCTTATCTAACCTCAACACTTCTTCCCTTTCATGCAATCTAATCTAC 241
DB 178 aacaactcttattctaaacctcaaacacttcttcccttcttcgaatcgaatcattacc 237

QY 242 CTCATCCCGGAGAAATGCGGAGTTCGGTATCAGAACGGTGCATCGGTTCTTTCA 301
DB 238 ctctatccgagagaaatgcccaggttccggtatcagaacgggtgactcggtttcttca 297

QY 302 CGATTATGAGTCTGCGAGATCATCGCATTCGGCTCATAAATACCTACTGCTTGGCC 361
DB 298 cgattatgagtcgtcgagatcatcgcattccgcctcaataactactgcttggccc 357

QY 362 CTAGCGAGATTATGAAGATCTCAGCTGCGCGCGATAGTCGAGGAATCTGATCGGAAT 421
DB 358 ctagcgagattatgaagatctcagctgcccgcgagatgcgagaaatctgatcggaat 417

QY 422 TGTTCAGATTTAGCCGTGCAAGTACTCTGTTCGGGTGCGGGGTGTCATFCGGA 481
DB 418 tgtttcaagttttagccgtgcaagtactctgttcccggttcgcggggtgctgcatcgga 477

QY 482 TAAGCAGCCACACACCGTTAGCGATGCTCCATGCTTTGTAGGTGTATGAAGATTTGGG 541
DB 478 taagcagccacacacccggttagcgatgctccatgcttctgtaggtttatgaagagttggg 537

QY 542 CTTCCATCACTAAATTCGGAGGAGATGATGAATTTTGGACGGAAGGTGAATGTTTCG 601
DB 538 cttccatcaactaaattcggagagatgatgaattcttggacggaagtgaaattgttcg 597

QY 602 CGGTTTTCGACCGATCGCTCGTAATATTCGCCCTAAATTCGACACATATTTATGGAA 661
DB 602 cggttttcgcacccgattcgctcgtaattatccgcttaatttcgacacatatttatggaa 657

DB 598 cggttttcgcacgcgctcgctgtaattatccgcctaaattggacacatatattatgaaca 657
QY 652 ACGGCGAGAAACGTCGCTTGAATCGCAGCATCCATCTTTTACCGCAGCATCGGATTCGAG 721
DB 658 acgcgcagaaacgctcgttggaatcgacatccttaccgacgagatcggttcgag 717
QY 722 CTACCTACTCTTTTACCCCAATCTGAAATTAAGAAATTAAGAGGTTTGAATTCAGAGAAAG 781
DB 718 ctactactctttcaccacaatctgaaattgaaggggttgattcagagaaag 777
QY 782 CCCCAATGTAGTTAATCTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 841
DB 778 ccccaatgtagttaact 837
QY 842 TCGCCAAATCGGTCGGAGATTACAAAGACGTGGATGACGACAAACGCGCTTCTTTTAA 901
DB 838 tcgccaatcggtcggagattacaaaagcgtggatgacgacaaacgcgttctctctctctctct 897
QY 902 TTCGATCGATTTAAGCGCGCTTTGGATCCCGCGCTTCCGGGGAACTTCTTCGGAAT 961
DB 898 ttccgatcgatttaagcgcgcttggatccgcggtcccggggaactactctcggaact 957
QY 962 GTCTATCGTTTTCGATGCGGAGATCTTGGCGCGGATTTGGTCCGAGATGAAGGGGTGT 1021
DB 958 gtctatcgtttgcgagtggaagatcctgcgcgggatttggctcgagatgaaggggtgt 1017
QY 1022 TTCGGGACGTCGAGCGATCGCGCGGAAATAGAGAGGACGACGACGACGACGACGACGACG 1081
DB 1018 ttccggcagctcgagcgatcgcgcggaataagagaagagagagagagagagagagagagag 1077
QY 1082 TAGAACTGTGAGAACGTGGCGCTGTGAGATTCGCGAAGCCTTCGCAAACTGTATTATTTCT 1141
DB 1078 tagaaactgtgagaactggcgtctgagattcgcgaagccttgcgaactgttattct 1137
QY 1142 CGGTGGCGGATCGAGCAGCTTGTATCTTACGGCGCGATTTGGATGGGTAGGCGG 1201
DB 1138 cgggtggcgagatcgagcaggttgcattcttcgcgcggttggatggggttaaggcg 1197
QY 1202 TGAAGCAAGAGATCTGTCGATGATGAGAGAGATTTACGATGCTGTGTGTAACCGA 1261
DB 1198 tgaagcaagagatcgtcgtatgagagagagagagagagagagagagagagagagagagag 1257
QY 1262 GGGATGCTGCGGAGGATTTGGAGGTGGATGCTTTCGCCAAAGAGGAAATTCGAAGGT 1321
DB 1258 gggatgctgcccggagattggaggttggtctcttccaaaggaggaattgcgaagctt 1317
QY 1322 TTGATGATTATTTTCGCGAGGAAATAAGGGTGTGATTAATCATTTAATCATGTATTATGA 1381
DB 1318 ttgatgattattttcggaggggaataaagggttgatttaattcaatcatgtattatga 1377
QY 1382 AGTTGGATGAATCCTCTGTTTCTCTATTGTTTAAACAATAATTTTTCATTTCCATTGAA 1441
DB 1378 agttggagcaaatcctctgttct 1437
QY 1442 CTTTTTTTCAGTCAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1501
DB 1438 cttttttgagccaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaaataaaaaa 1497
QY 1502 TTT 1522
DB 1498 ttt 1518

RESULT 2
AAT37311
ID AAT37311 standard; cDNA to mRNA; 1479 BP.
XX
AC AAT37311;
XX
DT 06-FEB-1997 (first entry)
XX
DE Aromatic acyl transferase coding sequence.
XX


```

Db 710 gaagaaactgctcaaatctcaggccacactactgttcttaaatccagcaatttctaaa 769
Qy 709 GATCGGATTCGAGCTACCTACCTTTTCACCCCAATCTGAATTAAGAAATGGAAGGTTTG 768
Db 770 gatgaagttcagccactctcctacacccctattgatcatatgaagctcagaataatc 829
Qy 769 ATTACAGAAAAGCCCCAAATGTA-----GTTAATCTCTCTTCCCTTCGTC 813
Db 830 atttcgtcaaaaaatcgcaacttaaccggtagtagtaattataatctgtcaactttcacg 889
Qy 814 GCGATCGGAGCTTATCTGACCGCCGATCGCCAAAT-----CGGTGCGAGATAC 864
Db 890 gfgacatctgcactgatctcggacatgcttgcgaataattagacacccgtcgaagag 949
Qy 865 AAAGACGTGGATGAGACAAACCGCTTTCTTTTAAATTCGATCCGATTAAGCGCGCT 924
Db 950 aagtggaagagataaaacatcagcaaatctatgtctttcatcaactgccgacaacgt 1009
Qy 925 TTGGATCCCGCGCTCCGGGGAACCTACTTCGGAACCTGCTTATCGTTTCGATGCGGAAG 984
Db 1010 ttgtcgcgcgatacctcaaatctacttggaaattgcatagtccttgatggtgga 1069
Qy 985 ATCTCGCGCGGATTTGGTCGAGATGAAGGGGTGTTTCGGCCAGCTGAGCGATCGCG 1044
Db 1070 tgcactcatgagcaactgttagaaatgaagggtgtgctgtagctgcaaccgcctcgga 1129
Qy 1045 GCGGAATAGACGAAGGAGCGACGACACAGAA---GATTCTAGAACTGTGAGAACTGG 1101
Db 1130 gatgctatccataaagggtttacatgactacagaaatctgagaggagattggtatcg 1189
Qy 1102 CCGTCTGAGATTCGCGAAGCCTTGCAAACT-----GTTATTCTCGGTGCGCGGA 1152
Db 1190 cgcgccgatacaatctcgcgcacaaaggtcgacgctcatattatgcttggatccgca 1249
Qy 1153 TCGAGAGCGCTTGATCTTTACCGCGGATTTTGAATCGGGTGAAGCGGTGAAGCAAGAG 1212
Db 1250 caacgcaatgcatgattttgtagcagattttggtggtggaagttgaaagcatgaa 1309
Qy 1213 ATACTGTCGATTCGATGAGAGAGTTTACGATGCTGTTGTGTAACCGAGGATGCTGCC 1272
Db 1310 tctgttcaactcaatctcctcggaacacataattttgatctcgtccgagaagatttaa 1369
Qy 1273 GGAGGATTCGAGTTGGATGCTTTGCGCAAGGAGGAATTCGAAGCTTTTGATGATTAT 1332
Db 1370 gtagcacttgacttggcaattcttctgctgaagaataggatggacgcatcttgcaccatt 1429
Qy 1333 TTTGCGGAGGAATAAAGGGT 1353
Db 1430 tttagcaatttcatcaatagt 1450

```

RESULT 6

AAF58252/c
ID AAF58252 standard; DNA; 936 BP.

XX AAF58252;

XX 24-APR-2001 (first entry)

XX Oligonucleotide D1835.

XX Electron-transfer group; ETM; mismatch; genotyping;
KW gene expression; ss.

XX Synthetic.

XX WO200107665-A2.

XX 01-FEB-2001.

XX 26-JUL-2000; 2000WO-US20476.

XX

PR 26-JUL-1999; 99US-0145695.
PR 17-MAR-2000; 2000US-0190259.
XX (CLIN-) CLINICAL MICRO SENSORS INC.
XX Umek RM;
XX WPI; 2001-159728/16.
XX Nucleic acids containiny electron-transfer group, useful as labels in hybridization assays, e.g. for genotyping, allowing repeat analyses on a single surface
XX Example 6; Page 127; 159pp; English.
XX The present invention relates to a composition comprising two nucleic acids each containing an electron-transfer group (ETM) having different redox potentials. The invention is used for electronic detection of nucleic acids, especially of substitutions (mismatches) and single-nucleotide polymorphisms, e.g. for genotyping, monitoring gene expression.
XX Sequence 936 BP; 4 A; 139 C; 10 G; 7 T; 776 other;

Query Match 5.7%; Score 87; DB 22; Length 936;
Best Local Similarity 6.2%; Pred. No. 4.3e-08;
Matches 33; Conservative 295; Mismatches 205; Indels 0; Gaps 0;

Qy 990 GCGGGCGGATTTGGTCGGAGATGAAGGGGTGTTTCGGGCAGCTGAGCGATCGCGCGGA 1049
Db 844 GGG 785
Qy 1050 AATAGAGAAGAGGACGACGACAAAGAGATTCTAGAAACTGTGGAACACTGCGCCGTCTGA 1109
Db 784 WWW 725
Qy 1110 GATTCGCGAAGCCTTGCAAACTGTTATTCTCGGTGCGGGATCGACGAGCTTCATCT 1169
Db 724 WWW 665
Qy 1170 TTACGCGCGGATTTTGGATGGGTGAAGCGGTGAAGCAAGAGATACTGCGATGATGG 1229
Db 664 WWW 605
Qy 1230 AGAAGAAGTTTACGATGCTGCTGTAAACCGAGGAGTGTGCGGAGGATTTGAGGTTGG 1289
Db 604 WWW 545
Qy 1290 ATTGTCTTTGCCAAAGGAGGAATTCGAAGCTTTTGATGATTATTTTTCGCGGGAATAA 1349
Db 544 WWW 485
Qy 1350 GGGTTGATTAAATCAATTAATCATGATTAATGAAGTTGGATGAATCCTCTGTTTCATCTC 1409
Db 484 WWW 425
Qy 1410 TATGTTTAAACAATAATTTTTCATTCGAACCTTTTTCAGTCAATAAAAAAATAA 1469
Db 424 WWW 365
Qy 1470 AAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 1522
Db 364 WWW 312

RESULT 7

AAF58254/c
ID AAF58254 standard; DNA; 936 BP.

XX AAF58254;

XX 24-APR-2001 (first entry)

DT	24-APR-2001	(first entry)
XX	Oligonucleotide D1954.	
XX	Electron-transfer group; ETM; mismatch; genotyping;	
KW	gene expression; ss.	
KW	Synthetic.	
OS	WO200107665-A2.	
PN	01-FEB-2001.	
PD	26-JUL-2000; 2000WO-US20476.	
XX	26-JUL-1999; 99US-0145695.	
PF	17-MAR-2000; 2000US-0190259.	
PR	(CLIN-) CLINICAL MICRO SENSORS INC.	
PR	Umek RM;	
PA	WPI; 2001-159728/16.	
XX	Nucleic acids containing electron-transfer group, useful as labels in	
PI	hybridization assays, e.g. for genotyping, allowing repeat analyses on	
DR	a single surface -	
XX	Example 6; Page 127; 159pp; English.	
PT	The present invention relates to a composition comprising two nucleic	
PT	acids each containing an electron-transfer group (ETM) having	
PT	different redox potentials. The invention is used for electronic	
PT	detection of nucleic acids, especially of substitutions (mismatches)	
PS	and single-nucleotide polymorphisms, e.g. for genotyping,	
XX	monitoring gene expression.	
SS	Sequence 936 BP; 5 A; 142 C; 7 G; 6 T; 776 other;	
CC	Query Match	5.3%; Score 80.2; DB 22; Length 936;
CC	Best Local Similarity	1.3%; Pred. No. 8.6e-07;
CC	Matches	4; Conservative 219; Mismatches 92; Indels 0; Gaps
QY	1208 AAGAGATACTGTCGATTGGATGTGAGAGAAGATTACGATGCTGTGTGTAAACCGAGGATG	1267
Db	249 WWWWWW	308
QY	1268 CTCGCCGAGGATTGGAGTTGGATTGCTTTGCCAAAGGGAATTGCAGCTTTTGATG	1327
Db	309 WWWWWW	368
QY	1328 ATTATTTTGGGAGGAATTAAGGGTTGATTAATCATTTAATCATGATTAATGAAGTTGG	1387
Db	369 WWWWWW	428
QY	1388 ATGAAATCCTCTCTCTCACTCTCTGTTGTTTAAACAATAATTTTTTCCATTGAACTTTT	1447
Db	429 WWWWWW	488
QY	1448 TGAGTCAATAAAAAAAAAAAAAAATGAAAAACTCAGTTATTTTTTTTTT	1507
Db	489 WWWWWW	548
QY	1508 TTTTTTTTTTTTTT 1522	
Db	549 WWWWWW	

[illegible]

Job time: 15463 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:04:56 ; Search time 168.74 seconds
(without alignments)
1707.547 Million cell updates/sec

Title: US-08-894-356C-6

Perfect score: 1522

Sequence: 1 NTGACCACCTCTCGAATC.....TTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 94655562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued_Patents_NA:*

1: /cgn2_6/ptodata/2/lna/5A_COMB.seq:*

2: /cgn2_6/ptodata/2/lna/5B_COMB.seq:*

3: /cgn2_6/ptodata/2/lna/6A_COMB.seq:*

4: /cgn2_6/ptodata/2/lna/6B_COMB.seq:*

5: /cgn2_6/ptodata/2/lna/PCTUS_COMB.seq:*

6: /cgn2_6/ptodata/2/lna/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	57	3.7	7218	1	US-08-232-463-14 Sequence 14, Appl
2	56.8	3.7	7218	1	US-08-232-463-14 Sequence 14, Appl
3	52.6	3.5	1875	5	PCT-US96-10618-1 Sequence 1, Appl
4	51.4	3.4	3138	1	US-07-867-106-4 Sequence 4, Appl
5	50	3.3	5173	1	US-08-242-677-1 Sequence 1, Appl
6	49.8	3.3	130	6	5198345-15 Patent No. 5198345
7	49.8	3.3	130	6	5198345-15 Patent No. 5198345
8	49.6	3.3	3138	1	US-07-867-106-4 Sequence 4, Appl
9	49.6	3.3	5852	1	US-07-867-106-2 Sequence 2, Appl
10	49.2	3.2	1578	4	US-09-416-050A-1 Sequence 1, Appl
11	49.2	3.2	1578	4	US-09-664-800-1 Sequence 1, Appl
12	49.2	3.2	1578	4	US-09-665-309-1 Sequence 1, Appl
13	49.2	3.2	1578	4	US-09-661-569-1 Sequence 1, Appl
14	48.4	3.2	1493	6	5340934-5 Patent No. 5340934
15	47.6	3.1	198	5	US-08-330-108-16 Sequence 16, Appl
16	47.6	3.1	198	5	PCT-US92-10087-16 Sequence 16, Appl
17	47.2	3.1	5852	1	US-07-867-106-2 Sequence 2, Appl
18	46.8	3.1	860	1	US-07-847-010-18 Sequence 18, Appl
19	46.8	3.1	4818	3	US-08-817-928-27 Sequence 27, Appl
20	46.6	3.1	1582	3	US-08-545-196B-10 Sequence 10, Appl
21	46.6	3.1	1582	3	US-08-545-196B-12 Sequence 12, Appl
22	46.6	3.1	22481	5	PCT-US95-07201-43 Sequence 43, Appl
23	46.4	3.0	2836	3	US-08-747-221B-24 Sequence 24, Appl
24	46.4	3.0	2836	3	US-08-747-221B-26 Sequence 26, Appl
25	45.6	3.0	3095	6	5231168-1 Patent No. 5231168
26	45.4	3.0	2209	1	US-08-514-014-1 Sequence 1, Appl
27	45.4	3.0	2209	2	US-08-833-823-1 Sequence 1, Appl

28 45.4 3.0 2280 3 US-08-813-150-1 Sequence 1, Appl
29 44.6 2.9 860 1 US-07-847-010-18 Sequence 18, Appl
30 44.6 2.9 2422 1 US-07-867-106-5 Sequence 5, Appl
31 43.8 2.9 222 4 US-08-481-190-15 Sequence 15, Appl
32 43.8 2.9 222 5 PCT-US93-00889-15 Sequence 15, Appl
c 33 43.6 2.9 83 1 US-08-420-443-5 Sequence 5, Appl
34 43.6 2.9 919 2 US-08-924-759-19 Sequence 19, Appl
35 43.6 2.9 919 3 US-09-248-335-19 Sequence 19, Appl
36 43.6 2.9 4507 2 US-08-568-459A-3 Sequence 3, Appl
37 43.6 2.9 4507 2 US-08-487-826B-3 Sequence 3, Appl
38 43.4 2.9 1931 3 US-09-019-942-2 Sequence 2, Appl
c 39 43.4 2.9 2502 3 US-09-234-332-1 Sequence 1, Appl
40 43.2 2.8 2058 2 US-08-749-391-1 Sequence 1, Appl
41 43.2 2.8 2058 3 US-09-390-200-1 Sequence 1, Appl
c 42 42.8 2.8 665 2 US-08-883-795A-36 Sequence 36, Appl
43 42.8 2.8 746 4 US-09-013-810-1 Sequence 1, Appl
44 42.8 2.8 1736 3 US-09-182-816-22 Sequence 22, Appl
c 45 42.8 2.8 1736 3 US-09-182-816-24 Sequence 24, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 23,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: pTZ9ptc-F1s
US-08-232-463-14

```
Query Match          3.7%; Score 57; DB 1; Length 7218;
Best Local Similarity 4.1%; Pred. No. 0.00037;
Matches 15; Conservative 210; Mismatches 140; Indels 0; Gaps 0;

QY 5 CCACCTCTCGAATCCTCCGAGTGGCGCGCTCCAGGACGCTGGCTGAGCAGTAC 64
Db 1099 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1158
QY 65 TCCCGCTCACCTCTCTGACATGAGCTGGCTGCAATTCACACCCATGCTTCAGCTTCTCT 124
Db 1159 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1218
QY 125 TCTAGAACTCCCTGTTCACAAACCGGCTCTCTCGAAACCGTCTCGAAACTCAAC 184
Db 1219 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1278
QY 185 ATCTCTATCTACCTCAACCTCAACACTTCTTCCCGCTTTCATGCAATCTAATCTACCCCTC 244
Db 1279 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1338
QY 245 TATCGCCGAGAAAAATGCGGAGTTCGCGATTCAGAACGGTGACTCGGTCTTCTTTCACGA 304
Db 1339 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1398
QY 305 TTATGGAGTCTGTGGAGATCATCGCATTCGCTGCAATAAATCTGCTTGGCCCTA 364
Db 1399 YYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYYY 1458
QY 365 GCGAC 369
Db 1459 ACTAC 1463

RESULT 2
US-08-232-463-14/c
; Sequence 14, Application US/08232463
; Patent No. 5670367
; GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHEIFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232.463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)836-9300
; TELEFAX: (703)683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
```

```
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZgpt-Fls
US-08-232-463-14

Query Match          3.7%; Score 56.8; DB 1; Length 7218;
Best Local Similarity 4.0%; Pred. No. 0.00041;
Matches 16; Conservative 225; Mismatches 157; Indels 0; Gaps 0;

QY 967 TCGTTTGGGATGGCGAAGATCCTGCGCGGGATTTGGTCGGAGATGAAGGGGTGTTTCGG 1026
Db 1462 TAGTTAAAGATAGAAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRR 1403
QY 1027 GCAGCTGAGGCGATCGCGCGGAAATAGAGAAGAGACGACGACAAAGAATTCAGAA 1086
Db 1402 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1343
QY 1087 ACTGTGGAGAACTGCCCTCTGAGATTCGGAAGCCTTGCAAACTGTATTCTCGGTG 1146
Db 1342 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1283
QY 1147 GCGGATCGACGAGCTTGATCTTTACGCGCGGATTTTGGATGGGTAAAGCGGTGAAG 1206
Db 1282 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1223
QY 1207 CAAGAGATACTGCAATTCGAGAGAGAAGTTTACGATGTCGTGTGTAACCGAGGAT 1266
Db 1222 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1163
QY 1267 GCTGCCGAGGATTCGAGGTGGATGTCCTTGGCAAGGAGGATTCGAAGCTTTTGTAT 1326
Db 1162 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1103
QY 1327 GATTATTTTGGGAGGAAATAAGGTTGATTATCAT 1364
Db 1102 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRAT 1065

RESULT 3
PCT-US96-10618-1
; Sequence 1, Application PC/TUS9610618
; GENERAL INFORMATION:
; APPLICANT: Coleman, Roger
; APPLICANT: Guegler, Karl J.
; APPLICANT: Au-Young, Janice
; APPLICANT: Bandman, Olga
; APPLICANT: Seilhamer, Jeffrey J.
; TITLE OF INVENTION: A NOVEL HUMAN EDG-2 RECEPTOR HOMOLOG
; NUMBER OF SEQUENCES: 6
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq Version 1.5
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US96/10618
; FILING DATE: 20-JUN-1996
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/000,352
; FILING DATE: 20-JUN-1995
```



```
; APPLICANT: Slade, Martin B
; APPLICANT: Chang, Andy C M
; APPLICANT: Williams, Keith L
; TITLE OF INVENTION: Improved plasmid Vectors for Cellular
; TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526rls
; STREET: One Liberty Place 46th Floor
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/867,106
; FILING DATE: 19920625
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: AU PJ 7187
; APPLICATION NUMBER: PCT/AU90/00530
; FILING DATE: 02-NOV-1989
; ATTORNEY/AGENT INFORMATION:
; NAME: Feeney, Joanne Longo
; REGISTRATION NUMBER: 35,134
; REFERENCE/DOCKET NUMBER: RICE-0002
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5852 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2378..5038
;
; US-07-867-106-2

Query Match 3.3%; Score 49.6; DB 1; Length 5852;
Best Local Similarity 69.8%; Pred. No. 0.017;
Matches 67; Conservative 0; Mismatches 29; Indels 0; Gaps 0;

QY 1427 TTTTTCATTCAGCTTTTTCAGTCAATATAAAAAAAAAAAAAAAAAAATGAAA 1486
Db 2032 TTTTTCATTCAGCTTTTTCAGTCAATATAAAAAAAAAAAAAAAAAAATGAAA 1973

QY 1487 AACTCAGTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1522
Db 1972 AAAGTTGGTAAACTACATTAGTTTTTTTATAGTTTT 1937

RESULT 10
US-09-416-050A-1
; Sequence 1, Application US/09416050A
; Patent No. 6194559
; GENERAL INFORMATION:
; APPLICANT: KIM, Soo Young
; TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Factors
; FILE REFERENCE: 1942/42
; CURRENT APPLICATION NUMBER: US/09/416,050A
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 83
```

```
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-416-050A-1

Query Match 3.2%; Score 49.2; DB 4; Length 1578;
Best Local Similarity 67.6%; Pred. No. 0.014;
Matches 69; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1388 ATGAAATCCTCTCTTTTCATCTCTATTGTTTAAACAATAATTTTTTCCATTGAACCTTTT 1447
Db 1450 atgaattcataatttcataaacatagtataaacttcagatttagctatagagaagatg 1509

QY 1448 TGAGTCAATAAAAAAAAAAAAAAAAAAAAAAAAAAATGAAAAA 1489
Db 1510 tgactaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1551

RESULT 11
US-09-664-800-1
; Sequence 1, Application US/09664800
; Patent No. 6218527
; GENERAL INFORMATION:
; APPLICANT: KIM, Soo Young
; TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Fac
; FILE REFERENCE: 1942/42
; CURRENT APPLICATION NUMBER: US/09/664,800
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/416,050
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-664-800-1

Query Match 3.2%; Score 49.2; DB 4; Length 1578;
Best Local Similarity 67.6%; Pred. No. 0.014;
Matches 69; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 1388 ATGAAATCCTCTCTTTTCATCTCTATTGTTTAAACAATAATTTTTTCCATTGAACCTTTT 1447
Db 1450 atgaattcataatttcataaacatagtataaacttcagatttagctatagagaagatg 1509

QY 1448 TGAGTCAATAAAAAAAAAAAAAAAAAAAAAAAAAAATGAAAAA 1489
Db 1510 tgactaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 1551

RESULT 12
US-09-665-309-1
; Sequence 1, Application US/09665309
; Patent No. 6232461
; GENERAL INFORMATION:
; APPLICANT: KIM, Soo Young
; TITLE OF INVENTION: Abscisic Acid Responsive Element -Binding Transcription Fac
; FILE REFERENCE: 1942/42
; CURRENT APPLICATION NUMBER: US/09/665,309
; PRIOR FILING DATE: 2000-09-19
; PRIOR APPLICATION NUMBER: 09/416,050
; PRIOR FILING DATE: 1999-10-12
; NUMBER OF SEQ ID NOS: 83
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1578
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
```


THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:01:00 ; Search time 5816.79 Seconds
(without alignments)
2473.399 Million cell updates/sec

Title: US-08-894-356C-6
Perfect score: 1522
Sequence: 1 NTGACCACCCCTCCTCGAATC.....TTTTTTTTTTTTTTTTTTTT 1522

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues 20456230
Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estro1:*
- 83: em_estro2:*
- 84: em_estro3:*
- 85: em_estro4:*
- 86: em_estro5:*
- 87: em_estro6:*
- 88: em_estro7:*
- 89: em_estro8:*
- 90: em_estro9:*
- 91: em_estro10:*
- 92: em_estro11:*
- 93: em_estro12:*
- 94: em_estro13:*
- 95: em_estro14:*
- 96: em_estro15:*
- 97: em_estro16:*
- 98: em_estro17:*
- 99: em_estro18:*
- 100: em_estro19:*
- 101: em_estro20:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

1

```

Qy 564 AGATGATGAATTCCT 578
Db 544 AGATGAACAATTCCT 558

RESULT 2
BG599447 690 bp mRNA EST 12-APR-2001
LOCUS EST504342 cSTS Solanum tuberosum cDNA clone cSTS25122 5' sequence,
DEFINITION mRNA sequence.
ACCESSION BG599447
VERSION BG599447.1 GI:13616583
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 690)
Bouffri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
Generations of ESTs from sprouting potato eyes
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13P-R.

FEATURES
source
1..690
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTS25122"
/clone_lib="cSTS"
/tissue_type="sprouting eyes from tubers"
/dev_stage="12-14 weeks post harvest"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
taken from tubers. The tubers were incubated at 26C in the
dark for 2-3 weeks prior to sprouting. The eyes were
frozen in liquid nitrogen immediately upon removal from
tubers."
BASE COUNT 178 a 156 c 144 g 212 t
ORIGIN

Query Match 10.3%; Score 157.2; DB 155; Length 690;
Best Local Similarity 55.2%; Pred. No. 1.7e-22;
Matches 306; Conservative 0; Mismatches 248; Indels 0; Gaps 0;

Qy 25 CGAGTGGCGCCCTCCAGGACAGGTGGCTGAGAGTCACTCCGGCTCACTTCTTCGAC 84
Db 9 CAAGTGGCGCCACCTCCGGCGCGCAGCGAGGTGACACTCCCTCTACTTATTTTGAT 68
Qy 85 ATGAGTGGCTGCTATTCACCCCATGCTTCAGCTTCTCTTACGAACTCCCTGTTC 144
Db 69 CATGTTGGTTAGGGTTCACCGGTATAGCGGGATATATTCTACAAGCTCCCATTTCC 128
Qy 145 AAACCCGCTTCTCGAAACCGTCTGTCGGAACCTCAACAACTCTTATCTTAACCTC 204
Db 129 AAACCCGATTTCTGTCAAACATATTCTCTCTTAAATAATTCACCTCTCCCTCACTC 188
Qy 205 AAACACTTCTCCCTTTCATGCAATCTATCTACCTCTCTACCGCGGAGAAATGCCG 264
Db 189 AAACACTATAGCCCTTAGCTGGAAACGTTGCTGCTCAGTATACAAACGGATATCCT 248
Qy 265 GAGTTCGGGTATCAGACGGTGACTCGGTTTCTTTCAGATATTGGAGTCTGCGGAGAT 324
Db 249 GAGTTACATTATGTGACAGGAGATCTGTATCTGTACTTCTTCTGAGACATGATGAT 308
Qy 325 CATCCGCATTCGGCTCAATAATACTACTGTCTTTCGCCCTAGCGACGATTTATGAAGATCTC 384

Db 309 TTCAATTATCTCATTTGGTGACCATCCGCGTAATGCTAAGGATTTTATCACTTTTCTCT 368
Qy 385 CAGCTGCGCGCGATAGTCGAGGAATCTGATCGGAATTTTCAAGTTTTCAGTCGCGTCAA 444
Db 369 ACGTTAGGGGAACCTAAGGATGCACCCGGGTCCAGCTTAGCTTCGATTCATCAA 428
Qy 445 GTGACTCTGTTTCCCGGTGCGGGGTGTCATCGGAATTAAGCAGCACACACCGCTTAGC 504
Db 429 GTGACACTTTTCCGATCTTGGTGATCCATTTGTTTCACTAACCATCATGTTGTTGGT 488
Qy 505 GATGCTCCATCGTTTGTAGGGTTTATGAAGAGTTGGGCTTCCATCACTAAATTCGGAGA 564
Db 489 GATGGAGCTACCATAGTAGGTTTCATTAGGGCGTGGGCTCTACTCCTAAATTCGGTGA 548
Qy 565 GATGATGAATTCCT 578
Db 549 GATGAACAATTCCT 562

RESULT 3
AW221049 606 bp mRNA EST 07-DEC-1999
LOCUS EST297518 tomato fruit mature green, TAMU Lycopersicon esculentum
DEFINITION cDNA clone cLEF3f1, mRNA sequence.
ACCESSION AW221049
VERSION AW221049.1 GI:6532733
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 606)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato fruit tissue
Unpublished (1999)
Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES
source
1..606
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3f1"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"
/notes="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; cLEF - Fruit were tagged at the lcm stage and
harvested 3-5 days prior to ripening. Fruit were cut in
half to verify the seeds were indeed 'immature' and the
seeds and locules were discarded prior to freezing the
pericarp"
BASE COUNT 147 a 143 c 123 g 193 t
ORIGIN

Query Match 9.6%; Score 146.6; DB 113; Length 606;
Best Local Similarity 53.4%; Pred. No. 2.6e-20;
Matches 308; Conservative 0; Mismatches 289; Indels 0; Gaps 0;

Qy 2 TGACCACCCTCCTCGAATCTCCCGAGTGGCGCGCTCCAGGCACGGTGGCTGAGCAGT 61

```

```

Db 5 TGCCACCGGATGAGCAATGCAAGTTGCGCCACTCCCGCGCGCAACGAGGTGA 64
QY 62 CACTCCCGCTCACCCTTCGCGACATGACGTGGCTGCTATTCACCCCAATGCTTCAGCTTC 121
Db 65 TACTCCCTCTTACTTATTTGACCATGTTTGGTTAGGGTTTCGCCGTATGAGCGGATAT 124
QY 122 TCTTCTACGAACCTCCCTGTTCCAAACCGCGCTTCCTCGAAACCGTTCGTTCCGAAACTCA 181
Db 125 TATTTTACAAGCTCTCCATTTTCAACCCGATTTCTGTTCAAAACATATATTCCTCCTCTTA 184
QY 182 ACAAATCCTTATCTTAACCTCAACACATCTTCCTCCCTTTTCATGCAATCTTAATCTACC 241
Db 185 AAAATTCACCTCTCCCTCACTCTCAACACACTATACGCCCTTAGCCGGAACGTTGCTTGTC 244
QY 242 CTCTATCGCGGAGAAATGCGCGAGTTCGGGTATCAGAACGCTGACTCGGTTTCCTTCA 301
Db 245 CACTAGATCAAAACGGATATCTCGATGCTGAGTTATGACAGAGATCTGTGCTGTGTTA 304
QY 302 CGATTATGAGTCTGTGCGGAGATCATCCGCAATTCCTGCTCATAAATACTACTGCTTGCCC 361
Db 305 CTTTTTCGAGACTGATATGAATTTCAATATCTCATTTGTTGACCATCCCGTAAGGCTA 364
QY 362 CTAGCGACCATATATGAAGATCTCCAGCTGCCCGCGATAGTCGAGGAATCTGTACGGAAT 421
Db 365 AGCATTTTATCACTTGTTCCTTAAGTTAGGGGAACCTAAGGATGCACCGGGGTCCAAC 424
QY 422 TGTTCCTCAAGTTTTCAGCGTCAAGTACACTCTGTTCCCGGTCCGGGGTGTGATCGGAA 481
Db 425 TAGCCCGCTCTTAGGCATTCAGGTGACACTTTTTCGGAATCTTGGTGTATCCATTGGTT 484
QY 482 TAACGACGACACACACGTTAGCGATGCTCCATCTGTTTGTAGGGTTTATCAAGAGTTGG 541
Db 485 TCACTAACCATCATGTTGTTGGTATGGAGCTACTATAGCAGGGTTTCATTAAGCGGTGG 544
QY 542 CTTCCTCACTAAATTCGGAGGAGATGATGAATCTT 578
Db 545 CTCTACTCCCAAAATTCGGTGGACATGAACAATCTT 581

RESULT 4
BE922784
LOCUS
DEFINITION
EST426553 potato leaves and petioles Solanum tuberosum cDNA clone
cSTB2E2 5' sequence, mRNA sequence.
ACCESSION
BE922784
VERSION
BE922784.1 GI:10448860
KEYWORDS
EST.
SOURCE
potato.
ORGANISM
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 570)
van der Hoeven,R.S., Bezlerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning
C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
source
1..570
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="cSTB2E2"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"

```

/note="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2: XhoI. Tissue was supplied by Dr. Fry (Cornell University). Leaflets and petioles were isolated from 8 week old greenhouse grown plants. The plants were watered and fertilized freely. The tissue was immediately frozen in liquid nitrogen."

BASE COUNT 139 a 136 c 115 g 180 t
ORIGIN

Query Match 9.1%; Score 138.8; DB 142; Length 570;
Best Local Similarity 53.5%; Pred. No. 1e-18;
Matches 290; Conservative 0; Mismatches 252; Indels 0; Gaps 0;

QY 2 TGACCACCCCTCCCTCGAATCTCCGAGTGGCGCGCTCCAGGCACGCTGCTGAGCAGT 61
Db 25 TGGGTGCTTTGATAGAGCAATGCAAGTTGCGCCACTCCGCGGCGGCGAGCGGAGTGA 84
QY 62 CACTCCCGCTCAGCTTCTTCGACATGACGTGGGTGCTATTTCCACCCCATGCTTCAGCTTC 121
Db 85 CACTCCCTCTTACTTATTTTGATCATGTTTGGTTAGGGTTCCACCGTATACGGCGGATAT 144
QY 122 TCTTCTACGAACCTCCCGTGTTCAAACCGCTTCCTCGAAACCGTGTTCGGAACCTCA 181
Db 145 TATTTCTACAAAGCTCCCATTTCCAAACCGATTTCTGTTCAAAACATATATTCCTCCTCTTA 204
QY 182 ACAAATCCTTATCTTAACCTCAACACACTTCTTCGCCCTTTTCATGCAATCTAATCTACC 241
Db 205 AAAATTCACCTCTCCCTCACTCTCAACACACTATAGCCCTTTAGCTGGAACGTTGCTCGTC 264
QY 242 CTCTATCGCGGAGAAATTCGCGAGTTCGCGTATCAGAACGGTGACTCGGTTTCCTTCA 301
Db 265 CACTAGATCAAAACGGATATCTCGAGTTACATATGTGACAGAGATCTGTATCTCTTA 324
QY 302 CGATTATGAGTCTGTGCGGAGATCATCCGCAATTCGCGCTCATAAATACTACTGCTTGC 361
Db 325 CTTTTCTGAGACTGATATGATTTCAATTTATCTCATTTGGTGACCATCCGCGTAATGCTA 384
QY 362 CTAGCGACGATATGAAGATCTCCAGCTGCCGCGATAGTCGAGGAATCTGATCGGAAT 421
Db 385 AGATTTTATCACTTTGTTCTTACGTTAGGGGAACCTAAGGATGCACCGGGGTCCAAT 444
QY 422 TGTTCCTCAAGTTTTCAGCGTCAAGTACTCTGTTTCCCGGTCCGGGGTGTGCTCGGAA 481
Db 445 TAGTCGCGGCTTAGCCATCAAGTGACACTTTTTCGAAATCTTGGTGTATCCATTGTT 504
QY 482 TAACGACGACACACACGTTAGCGATGCTCCATGTTTGTAGGGTTTATGAAGAGTTGG 541
Db 505 TCACTAACCATCATGTTGTTGGTATGGAGCTACCATAGTAGGGTTTCATTAGGCGGTGG 564
QY 542 CT 543
Db 565 CT 566

RESULT 5
AW221050
LOCUS
DEFINITION
EST297519 tomato fruit mature green, TAMU Lycopersicon esculentum
cDNA clone cLEF3F3, mRNA sequence.
ACCESSION
AW221050
VERSION
AW221050.1 GI:6533734
KEYWORDS
EST.
SOURCE
tomato.
ORGANISM
Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
1 (bases 1 to 583)
Alcala,J., Vrebalov,J., White,R., Matern,A.L., Holt,I.E., Liang,F.,
Upton,J., Hansen,T., Craven,M.B., Bowman,C.L., Ahn,S., Ronning,C.M.,
Fraser,C.M., Martin,G.B., Tanksley,S.D. and Giovannoni,J.

TITLE Generation of ESTs from tomato fruit tissue
JOURNAL Unpublished (1999)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.

FEATURES

Location/Qualifiers
1. .583
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEF3F3"
/clone_lib="tomato fruit mature green, TAMU"
/tissue_type="fruit pericarp"
/dev_stage="mature green (3-5 days pre-ripening)"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; cLEF - Fruit were tagged at the 1cm stage and harvested 3-5 days prior to ripening. Fruit were cut in half to verify the seeds were indeed 'immature' and the seeds and locules were discarded prior to freezing the pericarp"

BASE COUNT 142 a 140 c 117 g 184 t
ORIGIN

Query Match 9.1%; Score 138.6; DB 113; Length 583;
Best Local Similarity 52.5%; Pred. No. 1.1e-18;
Matches 303; Conservative 0; Mismatches 274; Indels 0; Gaps 0;

Qy 2 TGACCACCTCTCGAATCTCCGAGTGGCGCCGCTCCAGGACGGTGGCTGAGCAGT 61
Db 5 TGGCCACCGTATTGAGCAATGTCAAGTTGGCCACCTCCGGCGGCGCAACGGAGTGA 64
Qy 62 CACTCCCGCTCACCTCTTCGACATGAGTGGTGCATTTCCACCCCATGCTTCAGCTTC 121
Db 65 TACTCCCTTACITATTTTGACCATGTTGGTTAGGTTTTCGGCGTATGAGCGGATAT 124
Qy 122 TCTTCTAGCAACTCCCTGTTTCCAAACCCGCTTCTCGAAACCGTCTCGGAAACTCA 181
Db 125 TATTTTACAAGCTCTCCATTTTCAAACCGATTTGTTCAAACCATTTATCTCCTCTTA 184
Qy 182 AACATCTTATCTACCTCAACACTTCTTCCCGCTTCATGCAATCTAATCTACC 241
Db 185 AAAATCACTCTCCCTCACTCTCAACACTATAGCCCTTAGCCGGAACGTTGCTTCTC 244
Qy 242 CTCTATCCCGGAGAAATGCGGAGTTCCGGTATCAGAACGGTCACTCGGTTCTTTCA 301
Db 245 CACTAGATACAACGGATATCTGATTTACGTTATGTGACAGGAGATTTCTGTCTGTTA 304
Qy 302 CGATPATGAGTCTGTGGAGATCATCCGATTCGGCTCATATAATPACTCTCTTGGCCC 361
Db 305 CTTTTTTCGAGACTGATATGAATTTCAATTTATCTCATTTGGTGACCATCCGCGTAAGGCTA 364
Qy 362 CTAGGACGATATGAAGATCTTCAGCTGCGCGCATAGTGCAGGAATCTCATCGGAAT 421
Db 365 AGGATTTTATCATTCTGTTCTTAAAGTTAGGGAACCTTAAGGATGCACCGGGGTCCAA 424
Qy 422 TGTTCAGATTTTATAGCCGTGAAGTACTCTGTTTCCCGGTGCGGGGTGTCATCGAA 481
Db 425 TAGCCCGCTCTTAGCCATTCAGGTGACACATTTTCCGAATCTTGGTGATCCATGGTT 484
Qy 482 TAACGACACACACACCTTACCGATGCTCCATGCTTTTGTAGGTTTATGAAGAGTTGGG 541
Db 485 TCACATAACCATCATGTTTCTTGGTGATGGAACACTATATAGCATGTTTCAATTAAGGCGTGG 544
Qy 542 CTTCCATCACTAAATTCGGAGAGATGATGATTCCT 578
Db 545 CTCCTACTCCCAAAATAGTGGACATGAACCAATCTT 581

RESULT 6

BE921494
LOCUS BE921494 564 bp mRNA EST 02-OCT-2000
DEFINITION EST425179 potato leaves and petioles Solanum tuberosum cDNA clone
CSTB13B18 5' sequence, mRNA sequence.
ACCESSION BE921494
VERSION BE921494.1 GI:10447486
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 564)
van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buell,C.R., Ronning,
C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
Generation of ESTs from potato leaves and petioles
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.

FEATURES

Location/Qualifiers
1. .564
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTB13B18"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"

/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; Tissue was supplied by Dr. Fry (Cornell University). Leaflets and petioles were isolated from 8 week old greenhouse grown plants. The plants were watered and fertilized freely. The tissue was immediately frozen in liquid nitrogen."
BASE COUNT 150 a 139 c 97 g 178 t
ORIGIN

Query Match 8.2%; Score 124.8; DB 142; Length 564;
Best Local Similarity 53.6%; Pred. No. 7.9e-16;
Matches 283; Conservative 0; Mismatches 242; Indels 3; Gaps 1;

Qy 28 GTGGCGCGCTCCAGGACCGTGGCTGAGCAGTCACTCCCGCTCACCTTCTTCGACATG 87
Db 40 GCGCACCTCCCGCGGCGGCGACAGCTAAGCTCCCTCTAATCTATTTGATCAT 99
Qy 88 ACGTGGCTGCATTTCCACCCCATGCTTCAGCTTCTCTTCTAGGAACCTCCCTGTTTCCAA 147
Db 100 ATATGGTTGGGTTGGCTATATGCGGATTTTATTTCTACAAGCTACCAATTTCCAAA 159
Qy 148 CCGCGCTTCTCGAACCCTGCTCCGAACTCAACAATCTTATCTCTAACCTCAAA 207
Db 160 CTCGATTTCTGTTCAACCATTTATCTACTCTTAAACATCTCACTCTCCTCACTCAAA 219
Qy 208 CACTTCTTCCCTTTCATGCAATCTAATCTACCTCTCTATCGCGGAGAAATCGCGAG 267
Db 220 CACTACACACCTTAGCTGGCAACATTCCTGTCTCCTCAATTCGAGTGGTTATCCTGAG 279
Qy 268 TTCCGGTATCAGAACGGTGAATCTGCTTTCTTCAGGATTTGAGTCTGTGGAGATCAT 327
Db 280 TTGGCTTATGTGACTGGAGATTCATATCTGTACTTTTATTGAGACTGATATGAATTC 339
Qy 328 CCGCATTCGGCTCATATAATCTACTGCTTTGCCCCCTACGAGGATTTATGAAGATCTCCAG 387
Db 340 AATCATCTCATTTGGTAACCATCTCTCGAAATG---CTAAGGATTTTATCCCATCTCTCAA 396

```

QY 388 CTGCCCGCCGATGTCGAGGAATCTGATCGGAATTTGTTCAAGTTTATGCGGTGCAAGTG 447
    || || || || || || || || || || || || || || || || || || || || || ||
Db 397 TTGGCACAACCTAAGGATGACCCGGGGTCAAACTAGTCCCGCTTAGCCATTCAAGTG 456
    || || || || || || || || || || || || || || || || || || || || || ||
QY 448 ACTCTGTTTCCCGGTGCGGGGTGTCGATCGGAATACGACGACACACCGTTAGCGAT 507
    || || || || || || || || || || || || || || || || || || || || || ||
Db 457 ACACCTTTTCCGAATCTTGCCATATCCGTTGGTGTAGTAACCATCATGTCGCTTGAT 516
    || || || || || || || || || || || || || || || || || || || || || ||
QY 508 GCTCCATCGTTTGTAGGGTTTATGAAGAGTTGGGCTTCCATCACTAAA 555
    || || || || || || || || || || || || || || || || || || || || || ||
Db 517 GGAATACCATCTGGAATTCATAAGAACATGGGGTTTACTCAACAAA 564
    || || || || || || || || || || || || || || || || || || || || || ||

RESULT 7
LOCUS BE922746 574 bp mRNA EST 02-OCT-2000
DEFINITION EST426515 potato leaves and petioles Solanum tuberosum cDNA clone
CSTB1L12 5' sequence, mRNA sequence.
ACCESSION BE922746
VERSION BE922746.1 GI:10448822
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 574)
AUTHORS van der Hoeven,R.S., Bezzerides,J., Holt,I.E., Liang,F., Cho,J.,
Utterback,T., Hansen,C.L., Doan,B., Bougri,O., Buehl,C.R., Ronning
,C.M., Fry,W.E., Tanksley,S.D. and Baker,B.
TITLE Generation of ESTs from potato leaves and petioles
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone request: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com.
FEATURES
source 1..574
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="CSTB1L12"
/clone_lib="potato leaves and petioles"
/tissue_type="leaflets and petioles"
/dev_stage="8 weeks old plants"
/lab_host="SOLR"
/notes="vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue was supplied by Dr. Fry (Cornell University).
Leaflets and petioles were isolated from 8 week old
greenhouse grown plants. The plants were watered and
fertilized freely. The tissue was immediately frozen in
liquid nitrogen."
BASE COUNT 162 a 123 c 106 g 183 t
ORIGIN

Query Match 8.1%; Score 123; DB 142; Length 574;
Best Local Similarity 54.9%; Pred. NO. 1.9e-15;
Matches 265; Conservative 0; Mismatches 215; Indels 3; Gaps 1;

QY 99 TTTTCCACCCCATGCTTCAGCTCTCTCTACGAACCTCCCTGTCCAAACCCGCTTCCT 158
    ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 3 TTTTGGCTATATGCGTCGGATTTATTTCTACAAGCTACCAATTTCCAAACTCGATTCGT 62
    ||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 159 CGAAACCGTCTGTCGGAACCTCAAAACATCTTATCTCTAACCCCTCAAAACATCTTCTCC 218
    |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 63 TCAAACCATTTATTCCTACTCTTAAACATTCACCTCCCTCACCTCTCAAACTACACACC 122
    |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 219 CCTTTCATGCAATCTAATCTACCGCTCTATCCCGGAGAAAATCGCGAGTTCGGTATCA 278
    || | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
Db 123 CTTAGCTGGGAACATTCCTTTGTCCACTAAATTCGAGTGGTTATCTCCTGAGTTGGTTATGT 182
    || | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |
QY 279 GAACGGTGACTCGGTTCTTTTTCAGGATTATGGAGTCTGTGCGGAGATCATCCGCAATCCGC 338
    || | |||| | |||| | |||| | |||| | |||| | |||| | |||| | |||| |

```

```

Db 183 GACTGGAGATCTATATCTGTTACTTTTATTGAGACTGATATGAATTTCAATCATCTCAT 242
    || || || || || || || || || || || || || || || || || || || || || ||
QY 339 TCATAAATACTACTGCTTTGCCCTCAGCAGATTTATGAAGATCTCCAGCTGCGCGCAT 398
    || || || || || || || || || || || || || || || || || || || || || ||
Db 243 TGTAAACCATCTCTCGAAATG---CTAAGGATTTTATCCCATCTCCTCAATTGGCACAAC 299
    || || || || || || || || || || || || || || || || || || || || || ||
QY 399 AGTCGAGGAATCTGATCGGAAATTTGTTTCAAGTTTATGCGGTGCAAGTGACTCTGTTTTC 458
    || || || || || || || || || || || || || || || || || || || || || ||
Db 300 TAAGGATGACCGGGGTAAACTAGTCCCGGTCTTAGCCATTCAGTGACACTTTTTC 359
    || || || || || || || || || || || || || || || || || || || || || ||
QY 459 CGTCCGCGGGTGTGCATCGGAATAACGACGACACCGTTAGCGATGCTCCATCGTT 518
    || || || || || || || || || || || || || || || || || || || || || ||
Db 360 GAATCTTGCCATATCCGTTGTTTGTAGTAACCATCATGTCGCTTGTGATGAAATAACCAT 419
    || || || || || || || || || || || || || || || || || || || || || ||
QY 519 TGTAGGGTTTATGAAGAGTTGGGCTTCCCATCACTAAATTCGGAGGAGATGATGAATCTTT 578
    || || || || || || || || || || || || || || || || || || || || || ||
Db 420 CGTGAATTCATATAAGAACATGCGGGTTTACTCAAAATTCGCGGGTGTGAACAGTGCCTT 479
    || || || || || || || || || || || || || || || || || || || || || ||
QY 579 GGA 581
    ||
Db 480 AGA 482
    ||

RESULT 8
LOCUS AW102336 687 bp mRNA EST 21-NOV-2000
DEFINITION sd86d06.y1 Gm-cl009 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
Gm-cl009-1284 5' similar to TR:Q9ZWB4 Q9ZWB4 F21M11.13 PROTEIN. ;,
mRNA sequence.
ACCESSION AW102336
VERSION AW102336.1 GI:6072949
KEYWORDS EST.
SOURCE soybean.
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; euosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
REFERENCE 1 (bases 1 to 687)
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelding,J., Coryell,V., Khanna
,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
,Y., Person,B., Svaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World
Parkway Circle St. Louis, Missouri 63134 For further information
call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314)
427-3324 or contact: clones@genomesystems.com or
info@genomesystems.com web site: www.genomesystems.com
Insert Length: 2097 Std Error: 0.00
Seq primer: -40RP from Gibco
High quality sequence stop: 431.
Location/Qualifiers
1..687
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl009-1284"
/clone_lib="Gm-cl009"
/lab_host="Xl10-Gold"
/notes="Vector: pBluescript II XR; Site_1: EcoRI; Site_2:
XhoI; The mRNA was isolated from entire roots of
2-month-old 'Williams' plants that were greenhouse grown
in 5-gallon pots. To suppress nodulation, Black Gold

```



```
QY 388 CTGCCCCGAGTACTCAGGAATCTGATCGGAATTTCTTTCAAGTTTACCGCGTGAAGTG 447
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 377 TTGGCACAACCTAAGGATGCACGGGGTCAAACTAGTCCCGCTCTAGCCATTTCAAGTG 436
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 448 ACTCTCTTTTCCGGTCCGGGTGTGCATCGGAATAACGACGACACACCGTTACCGAT 507
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 437 ACATTTTTCGGAATCTTGGCATATCCGTTGGTTTGTAGTAACCATCATGTCGCTTGAT 496
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 508 G 508
    |
Db 497 G 497

RESULT 10
LOCUS BE434257
DEFINITION EST405335 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG15H6, mRNA sequence.
ACCESSION BE434257
VERSION BE434257.1 GI:9432100
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 591)
AUTHORS Alcalá,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
,S.D.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.
FEATURES
    Location/Qualifiers
        1..591
            /organism="Lycopersicon esculentum"
            /cultivar="TA496"
            /db_xref="taxon:4081"
            /clone_lib="cLEG15H6"
            /clone_lib="tomato breaker fruit, TIGR"
            /tissue_type="Pericarp"
            /dev_stage="breaker"
            /lab_host="SOLR"
            /note="Vector: pBluescriptSKmCquadpt; Site_1: EcoRI;
            Site_2: XhoI; Fruit were harvested at the breaker stage
            (first sign of lycopene accumulation on the blossom end of
            the fruit). Fruit were cut in half and the seeds and
            locules were discarded prior to freezing the pericarp."
BASE COUNT 159 a 129 c 121 g 182 t
ORIGIN

Query Match 7.6%; Score 116; DB 167; Length 591;
Best Local Similarity 52.7%; Pred. No. 5.1e-14;
Matches 251; Conservative 0; Mismatches 225; Indels 0; Gaps 0;

QY 103 CACCCCATGCTTCAGTCTCTTCTACGAACTCCCTGTTCCAAACCGCGCTTCCTCGAA 162
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 1 CGCCGTATGAGCGGATATATTATTTACAAGCTCTCCATTTTCAAAACCGGATTTTCGTCAA 60
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 163 ACCGTGTTCCGAACTCAACAATCCTTATCTACCTCAACCTCAACACTTCTTCCCCCTT 222
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 61 AACATTATTCCTCTTAAAAAATTCACCTCTCCCTCACTCTCAACACTATACGCCCTTA 120
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
```

```
QY 223 TCATGCAATCTAATCTACCTCTATCGCCGGAGAAAATCCGGAGTTCGGGTATCAGAAC 282
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 121 GCCGAAACGTTGCTTGTCCACTAGATACAAACGGATATCTCTGAGTTAGTTATGTGACA 180
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 283 GGTGACTCGGTTCCTTTCACGATTATGGAGTCTGTGCGGAGATCATCCGATTCCTCGTCA 342
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 181 GGAGATTCGTGTCTGTACTTTTTCGAGACTGATATGAATTTCAATTTATCTCATTTGGT 240
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 343 AAATACTACTGCTTTTGCCTTAGCGACGATTATGAAGATCTCCAGCTCGCGCGGATAGTC 402
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 241 GACCATCCCGCTAAGCTAAGGATTTTATCACTTTGTTCTTAAGTTAGGGGAACCTAAG 300
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 403 GAGGAATCTGATCGGAAATGTTCAAAGTTTTCAGCGTCAAGTACTCTGTTTCCCGGT 462
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 301 GATGCACCGGGGTCCAACTAGCCCGCTCTTAGCCATTTCAGGTGACACTTTTTCGGAAT 360
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 463 CGCGGGTGTGCATCGGAATAACGACGACACACCGCTTAGCGATGCTCCATGCTTTGTA 522
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 361 CTTGGTCTATCCATTTGTTTCACTAACCATCATGTTGTTGGTGATGGAGCTACTATAGCA 420
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 523 GGGTTTATGAAGAGTTGGGTTCCATCACTAAATTCGAGGAGATGATGAATTTCTT 578
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 421 GGGTTCAATTAAGCGGTGGGCTCTACTCCACAAATTCGGTGGACATGAACAATTTCTT 476
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 11
LOCUS BG598154
DEFINITION EST496832 cSTS Solanum tuberosum cDNA clone cSTS20M23 5' sequence,
mRNA sequence.
ACCESSION BG598154
VERSION BG598154.1 GI:13616294
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE 1 (bases 1 to 659)
AUTHORS van der Hoeven,R., Bezzerides,J., Sun,H., Cho,J., Chiemingo,A.,
Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and Baker,B.
TITLE Generations of ESTs from sprouting potato eyes
JOURNAL Unpublished (2000)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cmaeresgen.com
Seq primer: M13F-R.
FEATURES
    Location/Qualifiers
        1..659
            /organism="Solanum tuberosum"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone_lib="cSTS20M23"
            /clone_lib="cSTS"
            /tissue_type="sprouting eyes from tubers"
            /dev_stage="12-14 weeks post harvest"
            /lab_host="SOLR"
            /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
            XhoI; Various sizes of sprouting eyes (2mm to 15mm) were
            taken from tubers. The tubers were incubated at 26C in the
            dark for 2-3 weeks prior to sprouting. The eyes were
            frozen in liquid nitrogen immediately upon removal from
            tubers."
BASE COUNT 178 a 163 c 112 g 206 t
ORIGIN

Query Match 7.5%; Score 114.8; DB 155; Length 659;
Best Local Similarity 52.9%; Pred. No. 9e-14;
Matches 293; Conservative 0; Mismatches 257; Indels 4; Gaps 2;
```



```
Qy 28 GTGGGCGCCCTCCAGGCAACGGTGGCTGAGCAGTCACTCCCGCTCACTCTTCTTGACATG 87
Db 107 GCGGCACCTCCACCGCGCGCAACAGAGCTAACGCTCCCTCTAACTATTTTGGATCT- 165
Qy 88 ACGTGGCTGCAATTTCCACCCCATGCTTCAGCTTCTCTCTACGAACTCCCTCTGTCCAAA 147
Db 166 ATATGGTTGGCTTTTGGCTATATCGCTGGGATTTTATCTTCAAGCTACCAATTTCCAAA 225
Qy 148 CCGGCTTCCCTCGAAACCGTCTCGGAACTCAAACTATCTTCTTAACCTCAAA 207
Db 226 CTCGATTCGTTCAACACCATTTATCCCTACTCTTAACATTCACCTCTCCCTCACTCAAA 285
Qy 208 CACTTCTCCCTTTCATGCAATCTAATCTACCTCTATCCCGGAGAAAATCCCGAG 267
Db 286 CACTACACACCTTAGCTGGCAACATGCTTGTCCACTAAATTCGAGTGGTTATCCTGAG 345
Qy 268 TTCGGTATCAGACGGTGACTCGGTTCTTTCAGGATTTATGGAGTCTGCGGAGATCAT 327
Db 346 TTGGGTTATGTGACTGGAGATTTATATCTCTTATCTCTTATTTATTTAGACTGATATGAATTC 405
Qy 328 CCGATTCGGCTCATAAATACTACTGCTTTCGCCCTAGCGACGATTTATGAAGATCTCCAG 387
Db 406 AATCATCTCATTTGGTAACCATCTCGAATGCTAAGGATTTTATCCCTTCATTCCTCAA 465
Qy 388 CTGGCGCGGATGTCGAGGAATCTGATCGGAAATTTGTTTCAAGTTTATGCGGTCAAGTG 447
Db 466 TTGGCACAACTAAGGATGACCGGGGGTCAAACTAGTCCCGGCTTAGCCATTCAAGTG 525
Qy 448 ACTCTGTTTCCCGTCGCGGGGTGT---GCATCGGAATAACGACGCCACACCGTTAGC 504
Db 526 ACACCTTTTCCGAATCTTGGCATATCCCGTTGGTGTTTAGTAACCCATCATGTGCGCTGT 585
Qy 505 GATGCTCCATCGTTGTAGGGGTATATGAAGAGTTGGGCTTCCATCACTAAATTCGGAGGA 564
Db 586 GATGGAATACCATCGTGAATTCATAAGAACATGGGGTTTACTCAACAATTCGCGGT 645
Qy 565 GATGATGAATTCCT 578
Db 646 GATGAACAGTGTCT 659
```

RESULT 12

```
AW650280
LOCUS EST 518 bp mRNA 04-APR-2000
DEFINITION EST328734 tomato germinating seedlings, TAMU Lycopersicon
esculentum cDNA clone cLEI12F13 5', mRNA sequence.
ACCESSION AW650280
VERSION AW650280.1 GI:7411518
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 518)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Giovannoni,J.J., Martin,G.B. and Tanksley
,S.D.
```

TITLE Generation of ESTs from germinating tomato seed

```
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@CLEMSON.EDU
5 prime sequence.
```

FEATURES

```
source Location/Qualifiers
1..518 /organism="Lycopersicon esculentum"
```

```
/cultivar="TA96"
/db_xref="taxon:4081"
/clone="cLEI12F13"
/clone_lib="tomato germinating seedlings, TAMU"
/tissue_type="whole seedlings"
/dev_stage="7 days post imbibition"
/notes="Vector: pBlueScript SK(-); Site_1: EcoRI; Site_2:
XhoI; 7 days post imbibition on water-agar. Mixed stage
whole germinating seedlings from seed coat emergence up
to two centimeters in seeds not showing obvious signs of
germination were discarded."
```

```
BASE COUNT 134 a 119 c 100 g 165 t
ORIGIN
```

```
Query Match 7.4%; Score 112; DB 119; Length 518;
Best Local Similarity 51.8%; Pred. No. 3.4e-13;
Matches 253; Conservative 0; Mismatches 235; Indels 0; Gaps 0;
```

```
Qy 91 TGGCTGCATTTCCACCCCATGCTTCAGCTTCTCTTACGAACCTCCCGCTGTTCCAAACCC 150
Db 4 TGGTTAGGTTTCGCGGTATGAGCGGATATATTTACAGCTCTCCATATCAACCC 63
Qy 151 GCCTTCCTCGAAACCGTGGTTCGAAACTCAAACTATCTCTTAACCTCAACAC 210
Db 64 GATTTCTCTCAAAACATATTTCTCTCTCTTAAAAAATTCACCTCTACCTCAACAC 123
Qy 211 TTCTTCCCTTTCATCAATCTAATCTACCTCTATCGCGGAGAAAATGCCGAGTTC 270
Db 124 TATACGCCCTTAGCCGGAAGGTTGTTGCCACATAGATAACACGGATATCTCGAGTTA 183
Qy 271 CGGTATCAGAACGGTGACTCGGTTCTTTTTCACGATATGAGAGTCTGTGAGAGTATCCCG 330
Db 184 CTTTATGTGACAGAGATGCTGAGTCTGTACTTTTTCGAGACTGATATCAATCTCAAT 243
Qy 331 CATTCGCTCATAAATACTACTGCTTTGCGCCCTAGCGAGCATATGAAGATCTCCAGTG 390
Db 244 TATCTCATTTGGTGACCATCCGCGTAAGGATTTTGTATCACTTTGTCTCTAAGTTA 303
Qy 391 CCGCGGATAGTCGAGGAATCTGATCGGAAATTTTCAAGTTTTCAGCGTTCGAAGTGACT 450
Db 304 GGGAACTTACGATGACCCGGGGTCCAACTAGCCCGCTCTTAGCCATTCAGGTGACA 363
Qy 451 CTGTTTCCCGTCCGGGGGTGTGATCGGAATAACGACGACACACCGTTAGCGATGCT 510
Db 364 CTATCTCCGAATCTTGTGTTATCCATTTGATTTCACTAACCATCATGTTGTGGTGATGA 423
Qy 511 CCATCGTTTGTAGGGTTTATGAAGATTTGGGCTTCCATCACTAAATTCGGAGGAGATGAT 570
Db 424 GCTACTATAACAGGGTTTCATTAAAGCGTGGGCTCTACTCCACAAATTCGGTGGACATGA 483
Qy 571 GAATTCCT 578
Db 484 CAATTCCT 491
```

RESULT 13

```
AW616206
LOCUS EST 487 bp mRNA EST 24-MAR-2000
DEFINITION EST307245 L. hirsutum trichome, Cornell University Lycopersicon
hirsutum cDNA clone cLHT1D15 5', mRNA sequence.
ACCESSION AW616206
VERSION AW616206.1 GI:7322240
KEYWORDS EST.
SOURCE Lycopersicon hirsutum.
ORGANISM Lycopersicon hirsutum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 487)
AUTHORS van der Hoeven,R.S., Bezzerides,J.L., Matern,A.L., Holt,I.E., Liang
,F., Hansen,T., Craven,M.B., Bowman,C.L., Ronning,C.M., Nierman,W.,
```


<hr/>					
TITLE	Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley,S.D. Generation of ESTs from wild tomato (<i>Lycopersicon hirsutum</i>) trichomes				
JOURNAL	Unpublished (2000)				
COMMENT	Contact: David Frisch Clemson University Genomics Institute Clemson University 100 Jordan Hall, Clemson, SC 29634, USA Tel: 864 656 4366 Fax: 864 656 4293 Email: dfrische@CLEMSON.EDU 5 prime sequence.				
<hr/>					
FEATURES	Location/Qualifiers				
source	l..487 /organism="Lycopersicon hirsutum" /db_xref="taxon:62890" /clone="cLHTD15" /clone_lib="L_hirsutum trichome, Cornell University" /tissue_type="trichome" /dev_stage="mixed stages" /note="Leaves of various stages were shaken in liquid nitrogen, shearing off trichomes. This procedure yielded a mixture of cells which is highly enriched for trichome likely with minor contaminations of other types of leaf cells."				
BASE COUNT	116 a	117 c	91 g	163 t	
ORIGIN					
<hr/>					
Query Match	7.18; Score 108; DB 118; Length 487;				
Best Local Similarity	51.7%; Pred. No. 2.3e-12;				
Matches	246; Conservative	0;	Mismatches	230;	Indels 0; Gaps 0;
<hr/>					
Qy	33	GCCGCGTCCAGGCACGGTGCTGAGCATCACTCCCGCTCACCTTCTTCGACATGACGTG	92		
Db	1	GCCACTCCAGCGCGCAACGAGGTGGTACTTCCTCTTACTTATTATTGACCATGTTTG	60		
<hr/>					
Qy	93	GCTGCATTTCCACGCCCATGCTTCAGCTTCTCTTCTACGAATCCCCTGTTCCAAACCGC	152		
Db	61	GTTAGGGTTTCACTTATTAGCGGATATTATTCTACAAGCTCTCCAATTCCAAAATCCGA	120		
<hr/>					
Qy	153	CTTCCTCGAAGCGTCGTTCCGAACTCAACAATCCTTATCTCTAACCCCTCAACACTTT	212		
Db	121	TTTCGTTCAAACATATTTCCTCTCTTAAAAATTAACATTTTCCCTCACTCTCAACACTA	180		
<hr/>					
Qy	213	CTTCCCCCTTCATGCAATCTAATCTACCCCTCTATCGCGGAGAAATGCCGGAGTCCG	272		
Db	181	TATGCCCTTAGCCGGAACGTTGCTTGTCCTAGTATACAAACGGATATCTGAGTTAG	240		
<hr/>					
Qy	273	GTATCAGAACGGTGACTCGGTTCTTTTCACGATTATGGAGTCTGTCGGAGATCATCCGCA	332		
Db	241	TTATGTGACAGGAGATCTGTGTCTGTTAAATTTTTCTGAGACTGATATGGATTTCATTA	300		
<hr/>					
Qy	333	TTCCGCTCATAAATACTACTCTCTTTGCCCTTAGCGACGATTATGAAGATCTCCAGCTGCC	392		
Db	301	TCTCATTTGGCGACCATCCGCGCTAATGCTTAAGGATTTTTTATCACTTTGTTCTTAAGTTAGG	360		
<hr/>					
Qy	393	GCCGATAGTCAGGAATCTGATCGGAATTCGTTCAAGTTTTTACCGTTCGAAGTCACTCT	452		
Db	361	GGAACCTAAGGATGCACCCGGGTCCAATTAGCCCGGCTCTTAGCCATTCAAGTGACACT	420		
<hr/>					
Qy	453	GTTTCCCGGTCCGCGGTGTCTCGGAATAACCGACGCCACACCGTTTACGGCATG	508		
Db	421	TTTTCCGAATCTTGGTGTATCCATGGTTTTCACTAACCATCATGTTGTTGGTGATG	476		
<hr/>					
RESULT	14				
BF096979					
LOCUS	BF096979	482 bp	mRNA	EST	19-OCT-2000
DEFINITION	EST400516 tomato nutrient deficient roots Lycopersicon esculentum				
cdna clone	cLEW18020	5' sequence, mRNA sequence.			
ACCESSION	BF096979				
VERSION	BF096979.1	GI:10902689			

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 16:58:18 ; Search time 8904.87 Seconds
(without alignments)
29.529 Million cell updates/sec

Title: US-08-894-356C-22
Perfect score: 17
Sequence: 1 GAYTYGGNTGGGGNAA 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1344157 seqs, 7733874588 residues

Total number of hits satisfying chosen parameters: 2688314

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:*

- 1: gb_ba1:*
- 2: gb_ba2:*
- 3: gb_ba3:*
- 4: gb_in1:*
- 5: gb_in2:*
- 6: gb_in3:*
- 7: gb_om:*
- 8: gb_ov:*
- 9: gb_pat1:*
- 10: gb_pat2:*
- 11: gb_ph:*
- 12: gb_pl1:*
- 13: gb_pl2:*
- 14: gb_pl3:*
- 15: gb_pl4:*
- 16: em_ba1:*
- 17: em_ba2:*
- 18: em_fun:*
- 19: em_htgo_hum:*
- 20: em_htgo_inv:*
- 21: em_htgo_rod:*
- 22: em_htg_hum1:*
- 23: em_htg_hum2:*
- 24: em_htg_hum3:*
- 25: em_htg_hum4:*
- 26: em_htg_hum5:*
- 27: em_htg_hum6:*
- 28: em_htg_hum7:*
- 29: em_htg_hum8:*
- 30: em_htg_inv1:*
- 31: em_htg_inv2:*
- 32: em_htg_other:*
- 33: em_htg_rod:*
- 34: em_hum1:*
- 35: em_hum2:*
- 36: em_hum3:*
- 37: em_hum4:*
- 38: em_hum5:*
- 39: em_hum6:*
- 40: em_hum7:*
- 41: em_in:*
- 42: em_om:*
- 43: em_or:*

- 44: em_ov:*
- 45: em_pat:*
- 46: em_ph:*
- 47: em_pl:*
- 48: em_ro:*
- 49: em_sts:*
- 50: em_sy:*
- 51: em_un:*
- 52: em_vl:*
- 53: gb_sts1:*
- 54: gb_sts2:*
- 55: gb_sts3:*
- 56: gb_sy:*
- 57: gb_un:*
- 58: gb_vl1:*
- 59: gb_vl2:*
- 60: gb_htg1:*
- 61: gb_htg2:*
- 62: gb_htg3:*
- 63: gb_htg4:*
- 64: gb_htg5:*
- 65: gb_htg6:*
- 66: gb_htg7:*
- 67: gb_htg8:*
- 68: gb_htg9:*
- 69: gb_htg10:*
- 70: gb_htg11:*
- 71: gb_htg12:*
- 72: gb_htg13:*
- 73: gb_htg14:*
- 74: gb_htg15:*
- 75: gb_htg16:*
- 76: gb_htg17:*
- 77: gb_htg18:*
- 78: gb_htg19:*
- 79: gb_htg20:*
- 80: gb_htg21:*
- 81: gb_htg22:*
- 82: gb_htg23:*
- 83: gb_htg24:*
- 84: gb_htg25:*
- 85: gb_pr1:*
- 86: gb_pr2:*
- 87: gb_pr3:*
- 88: gb_pr4:*
- 89: gb_pr5:*
- 90: gb_pr6:*
- 91: gb_pr7:*
- 92: gb_pr8:*
- 93: gb_pr9:*
- 94: gb_rol:*
- 95: gb_rod:*
- 96: gb_in4:*
- 97: gb_pr10:*
- 98: em_ba3:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	14.2	83.5	850	13	AF227981
2	14.2	83.5	915	53	CNS06W8Q
3	14.2	83.5	1080	9	A16155
4	14.2	83.5	1080	9	A20628
5	14.2	83.5	1080	9	AR003699
6	14.2	83.5	1080	10	I28271
7	14.2	83.5	1096	9	AR110778
8	14.2	83.5	1320	13	AF053307
					AF227981 Euphorbia
					AL418128 T3 end of
					A16155 PTOM36 10/
					A20628 PTOM36 frag
					AR003699 Sequence
					I28271 Sequence 1
					AR110778 Sequence
					AF053307 Catharant

9	14.2	83.5	1388	13	AF297618	Taxus cus
10	14.2	83.5	1419	13	AF190130	Taxus cus
11	14.2	83.5	1436	9	AX025512	Sequence
12	14.2	83.5	1471	9	AX025518	Sequence
13	14.2	83.5	1476	12	AB029340	Perilla f
14	14.2	83.5	1479	10	E12756	Perilla ocl
15	14.2	83.5	1508	10	E12757	Senecio cru
16	14.2	83.5	1512	14	CPMPEL2	Z70521 C.melo mRNA
17	14.2	83.5	1526	9	A65937	Sequence 1
18	14.2	83.5	1542	10	I16758	Sequence 1
19	14.2	83.5	1563	14	NTHSR201	X95343 N.tabacum m
20	14.2	83.5	1605	10	E12755	Sequence 1
21	14.2	83.5	1605	12	AB026495	Petunia x
22	14.2	83.5	1622	10	E12754	Gentianatri
23	14.2	83.5	1622	12	AB026494	Gentiana
24	14.2	83.5	1679	12	AB010708	Gentiana
25	14.2	83.5	1703	10	E12753	Gentianatri
26	14.2	83.5	1735	14	DCHGBT3	Z84571 D.caryophyl
27	14.2	83.5	3706	10	I16766	Sequence 16
28	14.2	83.5	4746	8	AF228714	Ictalurus
29	14.2	83.5	8222	3	LLA250129	Lactococc
30	14.2	83.5	10238	3	U32728	Haemophilus
31	14.2	83.5	11148	1	AE006520	Streptococ
32	14.2	83.5	12850	2	AF282249	Lactococc
33	14.2	83.5	16078	65	AC019969	Drosophil
34	14.2	83.5	17500	64	AC017179	Drosophil
35	14.2	83.5	59888	62	AC011588	Homo sapi
36	14.2	83.5	66237	12	AB016892	Arabidops
37	14.2	83.5	69348	65	AC020525	Drosophil
38	14.2	83.5	71494	77	AC090191	Homo sapi
39	14.2	83.5	74342	12	AB020742	Arabidops
40	14.2	83.5	78174	4	AC004280	Drosophil
41	14.2	83.5	79676	12	AB013396	Arabidops
42	14.2	83.5	82360	13	AP000606	Arabidops
43	14.2	83.5	83650	12	AB023041	Arabidops
44	14.2	83.5	87581	14	ATT7H20	Arabidops
45	14.2	83.5	87841	12	AB005247	Arabidops
ALIGNMENTS						
RESULT 1						
AF227981	AF227981	850 bp	mRNA	PLN	17-FEB-2000	
LOCUS	Euphorbia esula F21J9.20-like protein mRNA, partial cds.					
DEFINITION	Euphorbia esula F21J9.20-like protein mRNA, partial cds.					
ACCESSION	AF227981					
VERSION	AF227981.1					
KEYWORDS	leafy spurge.					
SOURCE	Euphorbia esula					
ORGANISM	Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids 1; Malpighiales; Euphorbiaceae; Euphorbia.					
REFERENCE	1 (bases 1 to 850)					
AUTHORS	Anderson, J.V. and Horvath, D.P.					
TITLE	Identification of mRNAs expressed in underground adventitious buds of Euphorbia esula (leafy spurge)					
JOURNAL	Unpublished					
REFERENCE	2 (bases 1 to 850)					
AUTHORS	Anderson, J.V. and Horvath, D.P.					
TITLE	Direct Submission					
JOURNAL	Submitted (26-JAN-2000) Plant Science, USDA/ARS, 1605 Albrecht Blvd., Fargo, ND 58105, USA					
FEATURES	Location/Qualifiers					
source	1..850					
	/organism="Euphorbia esula"					
	/db_xref="taxon:3993"					
	/clone="12a"					
	/tissue_type="underground adventitious buds"					
	/dev_stage="3-day induced (defoliated)"					
	<1..850					
	/note="similar to Arabidopsis thaliana F21J9.20 encoded by					
CDS						

GenBank Accession Number AC000103"									
/codon_start=1									
/evidence=not_experimental									
/product="F21J9.20-like protein"									
/protein_id="AAF34801.1"									
/db_xref="GI:6984226"									
/translation="KRFVDFGSGKIGALKEIAASIAIPKPRVREAVSGLIWKGIITAFK									
KSNPLIRPSVMSVSNLRPTTPPVPAEAKLSKIDFEMSGWCREFGLYDAFCWG									
MODFVENYKVKQGEDGVCAICEFGKFAEKALSKIDFEMSGWCREFGLYDAFCWG									
KPTWLISIVSTNIRNVCILLDTKDGEGFEAWITLSEDSMWSFESDERVLEFAQVNPVGT									
L*									
3'UTR	661..850								
polyA_site	827								
BASE COUNT	258 a	123 c	227 g	242 t					
ORIGIN									
Query Match	83.5%;	Score 14.2;	DB 13;	Length 850;					
Best Local Similarity	76.5%;	Pred. No. 9.2e+02;							
Matches	13;	Conservative	2;	Mismatches	2;	Indels	0;	Gaps	0;
QY	1	GATTTGCGTGGGNA	17						
Db	466	GATTTGCGTGGGNA	482						
RESULT 2									
CNS06W8Q/c	915 bp	DNA	STS	11-JAN-2001					
LOCUS	T3 end of clone AX0AA029C10 of library AX0AA from strain CBS 7064								
DEFINITION	of Pichia farinosa, sequence tagged site.								
ACCESSION	AL418128								
VERSION	AL418128.1								
KEYWORDS	STS.								
SOURCE	Pichia farinosa.								
ORGANISM	Pichia farinosa								
REFERENCE	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Pichia.								
AUTHORS	1 (bases 1 to 915)								
TITLE	de Montigny, J., Spehner, C., Souciet, J., Tekala, F., Dujon, B., Wincker, P., Artiguenave, F. and Potier, S.								
JOURNAL	Genomic Exploration of the Hemiascomycetous Yeasts: 15. Pichia sorbitophila								
PUBMED	FEBS Lett. 487 (1), 87-90 (2000)								
AUTHORS	11152890								
TITLE	2 (bases 1 to 915)								
JOURNAL	Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G., Bolotin-Fukuhara, M., Bon, E., Brottier, P., Casaregola, S., de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B., Malpertuy, A., Neuveglise, C., Ozier-Kalogeropoulos, O., Potier, S., Saurin, W., Tekala, F., Toffano-Nioche, C., Wesolowski-Louvel, M., Wincker, P. and Weissenbach, J.								
PUBMED	Genomic Exploration of the Hemiascomycetous Yeasts: 1. A set of yeast species for molecular evolution studies(1)								
AUTHORS	FEBS Lett. 487 (1), 3-12 (2000)								
TITLE	11152876								
JOURNAL	3 (bases 1 to 915)								
PUBMED	Genoscope.								
AUTHORS	Direct Submission								
TITLE	Submitted (08-SEP-2000) Genoscope - Centre National de Sequencage, 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail : seqref@genoscope.cns.fr - Web : www.genoscope.cns.fr)								
JOURNAL	This STS is part of a random genomic sequencing program of thirteen yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces exiguus, Saccharomyces servazzii, Zygosaccharomyces rouxii, Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila, Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to 5 kb were prepared and both extremities were sequenced. See keywords for description of this sequence and for the sequence of the other extremity of this insert.								
COMMENT	Location/Qualifiers								
FEATURES									

source		1..915		/organism="Pichia farinosa"		/strain="CBS 7064"		/db_xref="taxon:4920"		/clone="AX0AA029C10"		/clone_lib="AX0AA"		/note="end : T3"	
STS		1..915		/note="end : T3"											
BASE COUNT		267 a		202 c		146 g		298 t		2 others					
ORIGIN															
Query Match		83.5%		Score 14.2;		DB 53;		Length 915;							
Best Local Similarity		76.5%		Pred. No. 9.1e+02;											
Matches 13;		Conservative 2;		Mismatches 2;		Indels 0;		Gaps 0;							
QY		1 GAYTTYGGNTGGGNAA 17													
Db		149 GACTTGGCTGGGAAA 133													
RESULT		3													
LOCUS		A16155		1080 bp		DNA		PAT		03-OCT-1994					
DEFINITION		PTOM36.													
ACCESSION		A16155													
VERSION		A16155.1		GI:640912											
KEYWORDS		synthetic construct.													
SOURCE		synthetic construct.													
ORGANISM		1 (bases 1 to 1080)													
REFERENCE		1													
AUTHORS		TRANSGENIC PLANTS WITH INCREASED SOLIDS CONTENT													
TITLE		Patent: WO 9314212-A 1 22-JUL-1993;													
JOURNAL		Location/Qualifiers													
FEATURES		1..1080													
source		/organism="synthetic construct"													
BASE COUNT		356 a		141 c		208 g		375 t							
ORIGIN															
Query Match		83.5%		Score 14.2;		DB 9;		Length 1080;							
Best Local Similarity		76.5%		Pred. No. 9e+02;											
Matches 13;		Conservative 2;		Mismatches 2;		Indels 0;		Gaps 0;							
QY		1 GAYTTYGGNTGGGNAA 17													
Db		866 GATTTGGATGGGAAA 882													
RESULT		4													
LOCUS		A20628		1080 bp		DNA		PAT		12-JUN-1994					
DEFINITION		PTOM36 fragment.													
ACCESSION		A20628													
VERSION		A20628.1		GI:583390											
KEYWORDS		synthetic construct.													
SOURCE		synthetic construct.													
ORGANISM		1 (bases 1 to 1080)													
REFERENCE		1													
AUTHORS		DNA, CONSTRUCTS, CELLS AND PLANTS DERIVED THEREFROM													
TITLE		Patent: WO 9105865-A 1 02-MAY-1991;													
JOURNAL		Location/Qualifiers													
FEATURES		1..1080													
source		/organism="synthetic construct"													
BASE COUNT		356 a		141 c		208 g		375 t							
ORIGIN															
Query Match		83.5%		Score 14.2;		DB 9;		Length 1080;							
Best Local Similarity		76.5%		Pred. No. 9e+02;											
Matches 13;		Conservative 2;		Mismatches 2;		Indels 0;		Gaps 0;							
QY		1 GAYTTYGGNTGGGNAA 17													
Db		866 GATTTGGATGGGAAA 882													
RESULT		5													
LOCUS		AR003699		1080 bp		DNA		PAT		04-DEC-1998					
DEFINITION		Sequence 1 from patent US 5744364.													
ACCESSION		AR003699													
VERSION		AR003699.1		GI:3964958											
KEYWORDS		Unknown.													
SOURCE		Unknown.													
ORGANISM		1 (bases 1 to 1080)													
REFERENCE		1													
AUTHORS		Bird,C.Roger, Grierson,D., Ray,J.Anthony and Schuch,W.Walter.													
TITLE		PTOM36 constructs and tomato cells transformed therewith													
JOURNAL		Patent: US 5744364-A 1 28-APR-1998;													
FEATURES		Location/Qualifiers													
source		1..1080													
BASE COUNT		356 a		141 c		208 g		375 t							
ORIGIN															
Query Match		83.5%		Score 14.2;		DB 9;		Length 1080;							
Best Local Similarity		76.5%		Pred. No. 9e+02;											
Matches 13;		Conservative 2;		Mismatches 2;		Indels 0;		Gaps 0;							
QY		1 GAYTTYGGNTGGGNAA 17													
Db		866 GATTTGGATGGGAAA 882													
RESULT		6													
LOCUS		I28271		1080 bp		DNA		PAT		06-FEB-1997					
DEFINITION		Sequence 1 from patent US 5569829.													
ACCESSION		I28271													
VERSION		I28271.1		GI:1819047											
KEYWORDS		Unknown.													
SOURCE		Unknown.													
ORGANISM		Unclassified.													
REFERENCE		1 (bases 1 to 1080)													
AUTHORS		Bird,C.R., Boniwell,J.M., Grierson,D., Ray,J.A. and Schuch,W.W.													
TITLE		Transformed tomato plants													
JOURNAL		Patent: US 5569829-A 1 29-OCT-1996;													
FEATURES		Location/Qualifiers													
source		1..1080													
BASE COUNT		356 a		141 c		208 g		375 t							
ORIGIN															
Query Match		83.5%		Score 14.2;		DB 10;		Length 1080;							
Best Local Similarity		76.5%		Pred. No. 9e+02;											
Matches 13;		Conservative 2;		Mismatches 2;		Indels 0;		Gaps 0;							
QY		1 GAYTTYGGNTGGGNAA 17													
Db		866 GATTTGGATGGGAAA 882													
RESULT		7													
LOCUS		AR110778		1096 bp		DNA		PAT		14-FEB-2001					

```
DEFINITION Sequence 1 from patent US 6127113.
ACCESSION AR110778
VERSION AR110778.1 GI:12827626
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1096)
AUTHORS Atkinson,R.L. and Dhurandhar,N.V.
TITLE Viral obesity methods and compositions
JOURNAL Patent: US 6127113-A 1 03-OCT-2000;
FEATURES
    source
        1. .1096
            Location/Qualifiers
                /organism="unknown"
BASE COUNT 353 a 215 c 227 g 301 t
ORIGIN

Query Match      83.5%; Score 14.2; DB 9; Length 1096;
Best Local Similarity 76.5%; Pred. No. 9e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGGNAA 17
   ||:|||||
Db 1013 GATTTGGATGGGTAA 1029

RESULT 8
AF053307
LOCUS AF053307 1320 bp DNA PLN 01-JAN-1999
DEFINITION Catharanthus roseus deacetylvindoline 4-O-acetyltransferase (DAT)
          gene, complete cds.
ACCESSION AF053307
VERSION AF053307.1 GI:4091807
KEYWORDS Madagascar periwinkle.
SOURCE Catharanthus roseus
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
          Magnoliophyta; eudicotyledons; core eudicots; Asteridae; euasterids
          I; Gentianales; Apocynaceae; Catharanthus.
REFERENCE 1 (bases 1 to 1320)
AUTHORS St-Pierre,B., Lafamme,P., Alarco,A.M. and De Luca,V.
TITLE The terminal O-acetyltransferase involved in vindoline biosynthesis
          defines a new class of proteins responsible for coenzyme
          A-dependent acyl transfer
JOURNAL Plant J. 14 (6), 703-713 (1998)
MEDLINE 98346012
REFERENCE 2 (bases 1 to 1320)
AUTHORS St-Pierre,B., Lafamme,P. and De Luca,V.
TITLE Direct Submission
JOURNAL Submitted (10-MAR-1998) Institut de Recherche en Biologie Vegetale,
          4101 Sherbrooke East, Montreal, QC H1X 2B2, Canada
FEATURES
    source
        1. .1320
            Location/Qualifiers
                /organism="Catharanthus roseus"
                /cultivar="Little Delicata"
                /db_xref="taxon:4058"
                <1. .>1320
                /gene="DAT"
                /product="deacetylvindoline 4-O-acetyltransferase"
                <1. .>1320
                /gene="DAT"
                1. .1320
                /gene="DAT"
                /EC_number="2.3.1.107"
                /function="involved in the last step in vindoline
                biosynthesis"
                /note="acetylcoenzyme A:deacetylvindoline
                17-O-acetyltransferase"
                /codon_start=1
                /product="deacetylvindoline 4-O-acetyltransferase"
                /protein_id="AAC99311.1"
                /db_xref="GI:4091808"

/translation="MESGKISVETETLSKTLIKPSSPTQSLSRYNLSYNDQNIYQTC
VSVGFYENPDGIEISTIREQLQNSLSTLVSYYPFAGKVKVKNNDYIHCNDDGIEFHV
RIRCRNMDILKELRSYVADVLVPRKVTGSEDTTAVOLSHFDCCGLAVAFGISHKV
ADGGTITASEMKDMAASACYLSSSHHVPTPLLVSDSIFPRQDNLICQFPTSKNCVEKT
FIFPPATEKLSKAVEFGIEKTRVEVLTAFLSRCATVAGKSAARNNCGQSLPPPV
LQAINLRPILLEPQNSVGNLSYIFSRITKENDYLNKEYTKLVINELRKEKQIKRNL
SREKLTYYVAQMEEFVKSLEFDISNFDLIDAYLSDSWCRFPFYDVFDMGKPIWVCILF
OPYIKNCVVMMDYPPFGDDYGEIAIVSFQEKMSAFEKNEQLQFVSN"

BASE COUNT 409 a 244 c 252 g 415 t
ORIGIN

Query Match      83.5%; Score 14.2; DB 12; Length 1320;
Best Local Similarity 76.5%; Pred. No. 8.8e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGGNAA 17
   ||:|||||
Db 1138 GATTTGGATGGGAAA 1154

RESULT 9
AF297618
LOCUS AF297618 1388 bp mRNA PLN 06-DEC-2000
DEFINITION Taxus cuspidata 2-debenzoyl-7,13-diacetylbaaccatin III-2-O-benzoyl
          transferase mRNA, complete cds.
ACCESSION AF297618
VERSION AF297618.1 GI:11559715
KEYWORDS Taxus cuspidata.
SOURCE Taxus cuspidata
ORGANISM Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
          Coniferopsida; Coniferales; Taxaceae; Taxus.
REFERENCE 1 (bases 1 to 1388)
AUTHORS Walker,K. and Croteau,R.
TITLE Taxol biosynthesis: Molecular cloning of a benzoyl- CoA:taxane
          2alpha -O-benzoyltransferase cDNA from Taxus and functional
          expression in Escherichia coli
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (25), 13591-13596 (2000)
MEDLINE 11095755
REFERENCE 2 (bases 1 to 1388)
AUTHORS Walker,K.D. and Croteau,R.B.
TITLE Direct Submission
JOURNAL Submitted (20-AUG-2000) Institute of Biological Chemistry,
          Washington State University, PO Box 646340, Pullman, WA 99164-6340,
          USA
FEATURES
    source
        1. .1388
            Location/Qualifiers
                /organism="Taxus cuspidata"
                /db_xref="taxon:99806"
                22. .1344
                /note="DBBT"
                /codon_start=1
                /product="2-debenzoyl-7,13-diacetylbaaccatin
                III-2-O-benzoyl transferase"
                /protein_id="AAG38049.1"
                /db_xref="GI:11559716"
                /translation="MGRFNVDMIERVIVAPCLQSPKNILHLSIDNKTGRLTNLSVY
                NASQSVSADPAKTREALSKVLVYPPFAGRLRNTENGDELEVEGAVEEAMA
                NDNLVLDQFNEDYDFPQQLVFNLRDENVIEDLHLLTVQTRFCGTVGFRFHSHV
                SDGKIGQLIKGMGEMARGFKPSLEPIWNRMVKPEDIMYLQDFHDFTHPPLNLEK
                SIQASNVISFERINYIKRCMMECKEKFSAFVVVALIWLAKSRFPDPNEVKYIKF
                PIDMRNFSPLPKGYGNACNACADNVKDLGSLLYALMLIKKSKFALNENKFS
                RILTKESTLDANMKHNENYCCGWRNLGFYADFNGNNAVNSPMQQRHELAMQNY
                FLFLSAKNMIDIGIKILMFMPASWVAPFKIEMEVTINKYVAKICNSKL"

BASE COUNT 446 a 243 c 301 g 398 t
ORIGIN

Query Match      83.5%; Score 14.2; DB 13; Length 1388;
Best Local Similarity 76.5%; Pred. No. 8.8e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
```


NIYPHNPSEGRDPVKVKEAIGKALVFFYPLAGRLREGGRKLFVECTGEGILFTEA
 DADVSLFEFWDITLFPYSSNNQNIHNAIENSDVNSPLLIQVTRKKGGFIFGLCF
 NHTMAGDFGVTQWPKATAETARGAFAPSLPVWQALLTARDPPRTFRHYEYDQVD
 MKSLIPVNSKIDQLEFESQLOISTLRQTLPAHLHDCPSPFVLTAVVRRLTALQFK
 PEEVREFLCVMNLRKSIDIPLGIVGNVAVPITTAACLGCPNPLGAYVDLIRKAKAK
 ATMEYIKSTVDLMVIRGPFYTVVGSFMSDLTRIGVENVDFGWRKAIFGGPTTGAR
 ITRGLVSFCVFPFMRNGERGTALSLCLPPAMERFRANVHASLQVQVQVDAVDSHMT
 IQSASK"

BASE COUNT 415 a 301 c 335 g 420 t
 ORIGIN

Query Match 83.5%; Score 14.2; DB 9; Length 1471;
 Best Local Similarity 76.5%; Pred. No. 8.7e+02;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 1123 GACTTTGGATGGGAAA 1139

RESULT 13

AB029340 1476 bp mRNA PLN 04-APR-2000
 LOCUS
 DEFINITION Perilla frutescens mRNA for anthocyanin acyltransferase, partial cds.

AB029340 1 GI:7415645
 ACCESSION
 VERSION
 KEYWORDS anthocyanin acyltransferase.
 SOURCE Perilla frutescens leaf cDNA to mRNA.
 ORGANISM Perilla frutescens

REFERENCE 1 (sites)
 AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
 Asteridae; euasterids I; Lamiales; Lamiaceae; Perilla.
 TITLE Sakakibara, K.Y., Tanaka, Y., Mizutani, M.F., Fujiwara, H., Fukui, Y.,
 Ashikari, T., Yamaguchi, M. and Kusumi, T.
 JOURNAL Molecular and biochemical characterization of a novel
 hydroxycinnamoyl-CoA: anthocyanin 3-O-glucoside-6
 unpublished (1999)

REFERENCE 2 (bases 1 to 1476)
 AUTHORS Sakakibara, K.Y. and Tanaka, Y.
 TITLE Direct Submission
 JOURNAL Submitted (24-JUN-1999) to the DDBJ/EMBL/GenBank databases. Keiko Y
 Sakakibara, Suntory Research Center, Fundamental Research, Plant
 Biotech; Shimamoto cho, Wakayamada, 1-1-1, Mishima-gun, Osaka
 618-8503, Japan (E-mail: Keiko_Sakakibara@suntory.co.jp,
 Tel: +81-75-962-8807, Fax: +81-75-962-8262)
 FEATURES
 Location/Qualifiers
 1..1476
 /organism="Perilla frutescens"
 /db_xref="taxon:48386"
 /tissue_type="leaf"
 <1..1343
 /codon_start=3
 /product="anthocyanin acyltransferase"
 /protein_id="BAA93475.1"
 /db_xref="GI:7415646"
 /translation="VETCRVGPDPDSVAEQSVPLTFDNTWLHFMPLQLLFYEPFC
 SKQHSIVPKLQSLKSTLIHFFPLNCLNLYPSSPERKMPFRLYLSGDSVSTIAES
 SDDFDLVGRVPLNLFVFKPLPIVEESDKLFQVFAVQVTLFPGRGVGIAT
 HHTVSDFSLFAPITAWSSMSKHIEDEDEEFKSLPFDVRSVYKPTKFDVIYMRNA
 LKPLQSRPLSPDTRIRTFVFTQSKIKKLGWIOSRPSVLSVLSLFSVAIAAMWAG
 ITKSFATDQDNEDEAFILPVDLPRDPDPENYFGNCLSVLPMSRMRRLVGVKGG
 VFLAAEVIAETKRIKINDRIETVEKWSPEPKALOKSYFSVAGSKSLDLVGADFGW
 GNARKEILSIDGETAMTLCRDEPGEGLVCLSLPKDKMDAFAAIFYSLGNG"

BASE COUNT 419 a 315 c 331 g 411 t

ORIGIN

Query Match 83.5%; Score 14.2; DB 12; Length 1476;
 Best Local Similarity 76.5%; Pred. No. 8.7e+02;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 1167 GATTTGGATGGGAA 1183

RESULT 14

E12756 1479 bp DNA PAT 24-JUN-1998
 LOCUS
 DEFINITION Perilla ocimoides mRNA for acyltransferase, partial cds.

E12756
 ACCESSION
 VERSION
 KEYWORDS E12756.1 GI:3251588
 SOURCE JP 1997070290-A/4.
 ORGANISM unidentified.
 unclassified.

REFERENCE 1 (bases 1 to 1479)
 AUTHORS Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y.,
 Yonekura, K., Mizutani, M. and Kusumi, T.
 TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
 JOURNAL Patent: JP 1997070290-A 4 18-MAR-1997;
 SUNTORY LTD

COMMENT OS Perilla ocimoides
 PN JP 1997070290-A/4
 PD 18-MAR-1997
 PF 30-JAN-1995 JP 1996046534
 PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
 ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
 MASAHIRO,
 PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAHI PC
 C12N15/09, A01H1/00, C07H21/04, C07K14/42, C12N9/10, (C12N9/10, PC
 C12R1:865),
 PC (C12N9/10, C12R1:19);
 CC strandedness: Double;
 CC topology: Linear;
 FH Key Location/Qualifiers
 FT source 1..1479
 /organism="Perilla ocimoides" FT
 /tissue_type="leaves"
 /clone="PSAT208"
 FT CDS 1..1343
 /product="acyltransferase".

FEATURES
 Location/Qualifiers
 1..1479
 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 420 a 316 c 331 g 411 t 1 others
 ORIGIN

Query Match 83.5%; Score 14.2; DB 10; Length 1479;
 Best Local Similarity 76.5%; Pred. No. 8.7e+02;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 1167 GATTTGGATGGGAA 1183

RESULT 15

E12757 1508 bp DNA PAT 24-JUN-1998
 LOCUS
 DEFINITION Senecio cruentus mRNA for acyltransferase, partial cds.

E12757
 ACCESSION
 VERSION E12757.1 GI:3251589
 SOURCE JP 1997070290-A/5.
 ORGANISM unidentified.
 unclassified.

REFERENCE 1 (bases 1 to 1508)
 AUTHORS Ashikari, T., Tanaka, Y., Fujiwara, H., Nakao, M., Fukui, Y.,
 Yonekura, K., Mizutani, M. and Kusumi, T.

TITLE GENE CODING PROTEIN HAVING ACYL GROUP TRANSFER ACTIVITY
JOURNAL Patent: JP 1997070290-A 5 18-MAR-1997;
SUNTORY LTD
COMMENT OS Senecio cruentus
PN JP 1997070290-A/5
PD 18-MAR-1997
PF 30-JAN-1996 JP 1996046534
PR 17-FEB-1995 JP 95P 67159, 29-JUN-1995 JP 95P 196915 PI
ASHIKARI TOSHIHIKO, TANAKA YOSHIKAZU, FUJIWARA HIROYUKI, PI NAKAO
MASAHIRO.
PI FUKUI YUKO, YONEKURA KEIKO, MIZUTANI MASAKO, KUSUMI TAKAAKI PC
C12N15/09,A01H1/00,C07H21/04,C07K14/42,C12N9/10,C12N9/10, PC
C12R1:865),
PC (C12N9/10,C12R1:19);
CC strandedness: Double;
CC topology: Linear; Location/Qualifiers
FH Key
FH
FT source 1..1508
FT /organism='Senecio cruentus'
FT /tissue_type='petal'
FT /clone='pCA48'
FT 1..1367
FT CDS /product='acyltransferase'.
FT Location/Qualifiers
1..1508
source /organism='unidentified'
/db_xref='taxon:32644'

BASE COUNT 442 a 293 c 296 g 477 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 10; Length 1508;
Best Local Similarity 76.5%; Pred. No. 8.7e+02;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;
Qy 1 GAYTTYGGNTGGGNAA 17
Db 1182 GATTTGGGTGGGGGAA 1198
|||||
|||||

Search completed: November 5, 2001, 16:58:20
Job time: 11270 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:12:08 ; Search time 445.49 Seconds
(without alignments)
23.961 Million cell updates/sec

Title: US-08-894-356c-22

Perfect score: 17

Sequence: 1 GAYTTYGGNTGGGNA 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Capext 1.0

Searched: 730101 seqs, 313950809 residues

Total number of hits satisfying chosen parameters: 1460202

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_0601.*
1: /SIDSI/gcgdata/geneseq/geneseq/NA1980.DAT.*
2: /SIDSI/gcgdata/geneseq/geneseq/NA1981.DAT.*
3: /SIDSI/gcgdata/geneseq/geneseq/NA1982.DAT.*
4: /SIDSI/gcgdata/geneseq/geneseq/NA1983.DAT.*
5: /SIDSI/gcgdata/geneseq/geneseq/NA1984.DAT.*
6: /SIDSI/gcgdata/geneseq/geneseq/NA1985.DAT.*
7: /SIDSI/gcgdata/geneseq/geneseq/NA1986.DAT.*
8: /SIDSI/gcgdata/geneseq/geneseq/NA1987.DAT.*
9: /SIDSI/gcgdata/geneseq/geneseq/NA1988.DAT.*
10: /SIDSI/gcgdata/geneseq/geneseq/NA1989.DAT.*
11: /SIDSI/gcgdata/geneseq/geneseq/NA1990.DAT.*
12: /SIDSI/gcgdata/geneseq/geneseq/NA1991.DAT.*
13: /SIDSI/gcgdata/geneseq/geneseq/NA1992.DAT.*
14: /SIDSI/gcgdata/geneseq/geneseq/NA1993.DAT.*
15: /SIDSI/gcgdata/geneseq/geneseq/NA1994.DAT.*
16: /SIDSI/gcgdata/geneseq/geneseq/NA1995.DAT.*
17: /SIDSI/gcgdata/geneseq/geneseq/NA1996.DAT.*
18: /SIDSI/gcgdata/geneseq/geneseq/NA1997.DAT.*
19: /SIDSI/gcgdata/geneseq/geneseq/NA1998.DAT.*
20: /SIDSI/gcgdata/geneseq/geneseq/NA1999.DAT.*
21: /SIDSI/gcgdata/geneseq/geneseq/NA2000.DAT.*
22: /SIDSI/gcgdata/geneseq/geneseq/NA2001.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	14.2	83.5	17	AAT37314	Degenerate primer
2	14.2	83.5	23	AAAI3989	Geranylgeranyl dip
3	14.2	83.5	1080	AAQ11802	Clone pTOM36. Lyc
4	14.2	83.5	1080	AAQ46682	PTOM36 cDNA clone.
5	14.2	83.5	1096	AAV07891	Nucleotide sequenc
6	14.2	83.5	1427	AAAC49764	Arabidopsis thalia
7	14.2	83.5	1429	AAC33965	Arabidopsis thalia
8	14.2	83.5	1436	AAC64784	Lemon alcohol acyl
9	14.2	83.5	1470	AAAC64787	Honey dew melon al
10	14.2	83.5	1477	AAC39128	Arabidopsis thalia
11	14.2	83.5	1479	AAT37311	Aromatic acyl tran

12	14.2	83.5	1508	17	AAT37312	Aromatic acyl tran	
13	14.2	83.5	1518	17	AAT37313	Aromatic acyl tran	
14	14.2	83.5	1526	18	AAT89415	Melon ripening-rel	
15	14.2	83.5	1542	15	AAQ54685	Anther specific cD	
16	14.2	83.5	1605	17	AAT37310	Aromatic acyl tran	
17	14.2	83.5	1622	17	AAT37309	Aromatic acyl tran	
18	14.2	83.5	1703	17	AAT37308	Aromatic acyl tran	
19	14.2	83.5	2865	21	AAC41615	Arabidopsis thalia	
20	14.2	83.5	2865	21	AAC46170	Arabidopsis thalia	
21	14.2	83.5	3706	15	AAQ54693	Anther specific ge	
22	14.2	83.5	3706	15	AAQ58340	Tobacco Ant32 geno	
23	13.2	77.6	300	20	AAQ12735	Human gene express	
24	13.2	77.6	311	16	AAT3671	Human gene signatu	
c	25	13.2	77.6	403	21	AAF15736	Human prostate can
c	26	13.2	77.6	432	21	AAC01828	Human secreted pro
27	13.2	77.6	652	21	AAZ80421	Human colon cancer	
c	28	13.2	77.6	680	18	AAZ30844	Streptococcus pneu
c	29	13.2	77.6	742	21	AAA01753	Human colon cancer
30	13.2	77.6	1155	21	AAZ53471	Neisseria meningit	
31	13.2	77.6	1379	21	AAC46852	Arabidopsis thalia	
32	13.2	77.6	1449	20	AAZ12045	Neisseria meningit	
33	13.2	77.6	1488	21	AAC64780	Strawberry vesca a	
34	13.2	77.6	1502	21	AAC49421	Arabidopsis thalia	
35	13.2	77.6	1504	21	AAC33536	Arabidopsis thalia	
36	13.2	77.6	1520	21	AAC42254	Arabidopsis thalia	
37	13.2	77.6	1564	20	AAZ56425	Clarkia breweri be	
38	13.2	77.6	1632	21	AAC64766	Strawberry alcohol	
39	13.2	77.6	1632	21	AAZ27666	Strawberry alcohol	
40	13.2	77.6	1648	21	AAC64785	Lemon acyl transfe	
41	13.2	77.6	1831	20	AAZ56427	Clarkia concinna b	
42	13.2	77.6	3007	20	AAZ56428	Clarkia concinna b	
43	13.2	77.6	3036	20	AAZ56426	Clarkia concinna b	
44	13.2	77.6	28473	19	AAZ52216	Streptococcus pneu	
45	13.2	77.6	92934	21	AAZ81473	N. meningitidis pa	

ALIGNMENTS

RESULT 1
AAT37314
ID AAT37314 standard; DNA; 17 BP.
XX
AC AAT37314;
XX
DT 06-FEB-1997 (first entry)
XX
DE Degenerate primer for identifying aromatic acyl transferase sequence.
XX
KW Aromatic acyl transferase; transformation; anthocyanin pigment;
KW plants; acylation; colour; tone; colouration; colour change;
KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
KW Scenecio cruentus; Lavandula angustifolia; ss.
XX
QS Synthetic.
XX
FH Key Location/Qualifiers
FT misc_feature 9 /*tag= a
FT /*mod_base= I
FT misc_feature 15 /*tag= b
FT /*mod_base= I
XX
XX WO9625500-A1.
XX
XX 22-AUG-1996.
XX
XX 16-FEB-1996; 96WO-JP00348.
XX
XX 30-JAN-1996; 96JP-0046534.
XX 17-FEB-1995; 95JP-0067159.
XX 29-JUN-1995; 95JP-0196915.

```

XX PA (SUNR ) SUNTORY LTD.
XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
XX PI Nakao M, Tanaka Y, Yonekura K;
XX XX WPI; 1996-393401/39.
XX DR DNA coding for aromatic acyl transferase - for transforming plants
XX PT which produce anthocyanin pigments and thus altering colour tone,
XX PT e.g. of flowers
XX XX
XX PS Claim 3; Page 82; 94pp; Japanese.
XX CC Vectors containing DNA fragments encoding proteins of plant origin
XX CC with aromatic acyl transferase activity may be used to transform
XX CC plants which produce anthocyanin pigments. The aromatic acyl
XX CC transferase acylates and allowing new colourations to be produced. Six
XX CC tone changes and allowing new colourations to be produced. Six
XX CC specific DNA sequences encoding aromatic acyl transferase from
XX CC different plants are described in AAT37308-T37313. This degenerate
XX CC primer was synthesised base on a peptide fragment (AAW04728)
XX CC isolated from three of the six clones. It was used to identify
XX CC other aromatic acyl transferase encoding clones.
XX XX
XX SQ Sequence 17 BP; 3 A; 0 C; 7 G; 3 T; 4 other;

Query Match 83.5%; Score 14.2; DB 17; Length 17;
Best Local Similarity 100.0%; Pred. No. 49;
Matches 17; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNA 17
Db 1 gaytvgntgggnaa 17

RESULT 2
AAA13989/C
ID AAA13989 standard; DNA; 23 BP.
XX AC AAA13989;
XX XX
XX DT 08-AUG-2000 (first entry)
XX DE
XX DE Geranylgeranyl diphosphate synthase PCR primer SEQ ID NO:6.
XX KW Geranylgeranyl diphosphate synthase; GGPP synthase; yew; cytostatic;
XX KW anticancer; Taxus; diterpene; pacitaxel; identification; plant;
XX KW Taxomyces andreanae; Penicillium raistrickii; microorganism;
XX KW PCR primer; ss.
XX OS
XX OS Taxus canadensis.
XX XX
XX PN US6043072-A.
XX XX
XX PD 28-MAR-2000.
XX XX
XX PF 05-NOV-1998; 98US-0187050.
XX XX
XX PR 05-NOV-1998; 98US-0187050.
XX XX
XX PA (UNIW ) UNIV WASHINGTON STATE RES FOUND.
XX XX
XX PI Croteau RB, Hefner JL;
XX XX
XX DR WPI; 2000-282526/24.
XX XX
XX PT Nucleic acid encoding geranylgeranyl diphosphate is useful for
XX PT producing pacitaxel and other diterpenes that are useful as anticancer
XX PT drugs
XX XX
XX PS Example 1; Column 39; 57pp; English.

```

```

XX CC The present sequence represents a PCR primer for a geranylgeranyl
XX CC diphosphate (GGPP) synthase protein. GGPP synthase has cytostatic
XX CC activity. A vector encoding GGPP synthase is useful in increasing
XX CC GGPP synthase levels in a host cell preferably taxus (yew) cell and
XX CC thereby facilitates production, isolation and purification of larger
XX CC amounts of GGPP synthase in plants. GGPP synthase is useful in obtaining
XX CC expression or enhanced expression of GGPP and other diterpenes, such as
XX CC pacitaxel, useful as anticancer drugs. Isolated nucleic acids encoding
XX CC GGPP synthase or hybridising with GGPP synthase encoding nucleic acids
XX CC are used for identifying genes encoding GGPP synthase from
XX CC microorganisms such as Taxomyces andreanae and Penicillium raistrickii.
XX XX
XX SQ Sequence 23 BP; 6 A; 9 C; 4 G; 4 T; 0 other;

Query Match 83.5%; Score 14.2; DB 21; Length 23;
Best Local Similarity 76.5%; Pred. No. 51;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNA 17
Db 20 GAYTTCGGCTGGGGTAA 4

RESULT 3
AAQ11802
ID AAQ11802 standard; DNA; 1080 BP.
XX AC AAQ11802;
XX XX
XX DT 23-JUL-1991 (first entry)
XX DE Clone pTOM36.
XX XX
XX KW Fruit ripening; pTOM36; ss.
XX OS
XX OS Lycopersicon esculentum.
XX PN WO9105865-A.
XX XX
XX PD 02-MAY-1991.
XX XX
XX PF 17-OCT-1990; 90WO-GB01603.
XX PR 20-OCT-1989; 89GB-0023716.
XX XX
XX PA (ICIL ) IMPERIAL CHEM INDS PLC.
XX XX
XX PI Bird CR, Grierson D, Ray JA, Schuch WW;
XX XX
XX DR WPI; 1991-148743/20.
XX XX
XX PT DNA constructs contg. DNA from pTOM36 clone - used to transform
XX PT plants to regulate prodn. of the fruit-ripening pTOM36 enzyme
XX XX
XX PS Disclosure; fig 1; 22pp; English.
XX XX
XX CC This clone is contained in a DNA construct used to transform host
XX CC plant cells for regulating the prodn. of the enzyme encoded by pTOM36.
XX CC Plants such as apple, tomato and mango may be modified to produce e.g.
XX CC sweeter fruit, novel flavour, modified colour or to have improved
XX CC processing characteristics. The DNA construct pref. also comprises
XX CC a constitutive or regulatory promoter, e.g. the CaMV 35S or poly-
XX CC galacturonase gene promoters.
XX XX
XX SQ Sequence 1080 BP; 357 A; 141 C; 207 G; 375 T; 0 other;

Query Match 83.5%; Score 14.2; DB 12; Length 1080;
Best Local Similarity 76.5%; Pred. No. 74;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 GAYTTYGGNTGGGNAA 17
 ||:|||| ||||| ||
 Db 866 gatttggatggggaaa 882

RESULT 4

AAQ46682
 ID AAQ46682 standard; cDNA; 1080 BP.

XX AC AAQ46682;

XX DT 18-JAN-1994 (first entry)

XX DE PTOM36 cDNA clone.

XX Transgenic fruit production; prodn.; increased solids content;
 KW higher reducing sugar content; tomato; tomatoes; plant cultivation;
 KW improved; paste; soup; ss.

XX OS Synthetic.

XX PN W09314212-A.

XX PD 22-JUL-1993.

XX PF 08-JAN-1993; 93WO-GB000021.

XX PR 10-JAN-1992; 92GB-0000520.

XX PA (ZONE) ZENECA LTD.

XX PI Bird CR, Boniwell JW;

XX DR WPI; 1993-243227/30.

XX Transgenic fruit with increased solids content - is obtd. by
 PT cultivating plants, esp. tomatoes, in which expression of genes
 PT homologous to PTOM36 is inhibited

XX PS Disclosure; Fig 1; 26pp; English.

XX The sequence is that of the cDNA clone PTOM36 which is believed to
 CC encode a cytoplasmic protein of approximately 52000 daltons
 CC involved in the ripening of tomatoes. The clone is 1069 bases long
 CC with an open reading frame of 271 codons. DNA homologous to PTOM36
 CC may be used in a construct that inhibits expression of genes
 CC homologous to PTOM36 during ripening. Fruit from plants transformed
 CC with this construct have increased solids content and a higher
 CC content of reducing sugars, e.g. glucose and fructose. The fruit
 CC may be mangoes, peaches, apples, pears, strawberries, bananas,
 CC melons and esp. tomatoes, which are useful in the prodn. of
 CC improved tomato paste and soup.

XX Sequence 1080 BP; 356 A; 141 C; 208 G; 375 T; 0 other;

Query Match 83.5%; Score 14.2; DB 14; Length 1080;
 Best Local Similarity 76.5%; Pred. No. 74;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNAA 17
 ||:|||| ||||| ||
 Db 866 gatttggatggggaaa 882

RESULT 5

AAV07891
 ID AAV07891 standard; cDNA; 1096 BP.

XX AC AAV07891;

XX DT 29-JAN-1999 (first entry)

XX

DE Nucleotide sequence of the fiber protein encoded by adenovirus Ad-36p.
 XX Adenovirus; Ad-36p fiber protein; viral obesity; vaccine; ss.
 XX Mastadenovirus.

XX FH Key Location/Qualifiers
 XX FT CDS 1..1096
 XX FT /*tag= a
 XX FT /product= "Ad-36p fiber protein"

XX PN W09844946-A1.

XX PD 15-OCT-1998.

XX PF 06-APR-1998; 98WO-US06730.

XX PR 04-APR-1997; 97US-004:942.

XX PA (OBET-) OBETECH LLC.

XX PI Atkinson RL, Dhurandhar NV;

XX DR WPI; 1998-568305/48.

XX Determining if obesity in a person is caused by Ad-36 virus - and
 PT providing the basis for treatment or prevention of obesity-causing,
 PT cholesterol reducing adenovirus, using the purified variant, Ad-36p

XX PS Claim 1; Pages 18-19; 24pp; English.

XX This is the nucleotide sequence encoding the adenovirus Ad-36p fiber
 CC protein used in the method of the invention to determine if a person
 CC is suffering viral obesity. The method is used to determine whether
 CC obesity in a person has a viral basis. Ad-36p can be used as a basis
 CC of a vaccine to prevent viral-based obesity.

XX Sequence 1096 BP; 353 A; 215 C; 227 G; 301 T; 0 other;

Query Match 83.5%; Score 14.2; DB 19; Length 1096;
 Best Local Similarity 76.5%; Pred. No. 74;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNAA 17

Db 1013 gatttggatggggtaa 1029

RESULT 6

AAC49764
 ID AAC49764 standard; DNA; 1427 BP.

XX AC AAC49764;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 62347.

XX Hybridisation assay; genetic mapping; gene expression control;
 KW protein identification; signal transduction pathway;
 KW metabolic pathway; promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX PR 25-FEB-1999; 99US-0121825.

XX PR 05-MAR-1999; 99US-01231E0.

PR 09-MAR-1999; 99US-0123548.
 PR 23-MAR-1999; 99US-0125788.
 PR 25-MAR-1999; 99US-0126264.
 PR 29-MAR-1999; 99US-0126785.
 PR 01-APR-1999; 99US-0127462.
 PR 06-APR-1999; 99US-0128234.
 PR 08-APR-1999; 99US-0128714.
 PR 16-APR-1999; 99US-0129845.
 PR 19-APR-1999; 99US-0130077.
 PR 21-APR-1999; 99US-0130449.
 PR 23-APR-1999; 99US-0130510.
 PR 28-APR-1999; 99US-0130891.
 PR 30-APR-1999; 99US-0131449.
 PR 30-APR-1999; 99US-0132048.
 PR 04-MAY-1999; 99US-0132407.
 PR 05-MAY-1999; 99US-0132484.
 PR 06-MAY-1999; 99US-0132485.
 PR 06-MAY-1999; 99US-0132486.
 PR 07-MAY-1999; 99US-0132487.
 PR 11-MAY-1999; 99US-0132863.
 PR 14-MAY-1999; 99US-0134256.
 PR 14-MAY-1999; 99US-0134218.
 PR 14-MAY-1999; 99US-0134219.
 PR 14-MAY-1999; 99US-0134221.
 PR 14-MAY-1999; 99US-0134370.
 PR 18-MAY-1999; 99US-0134768.
 PR 19-MAY-1999; 99US-0134941.
 PR 20-MAY-1999; 99US-0135124.
 PR 21-MAY-1999; 99US-0135353.
 PR 24-MAY-1999; 99US-0135629.
 PR 25-MAY-1999; 99US-0136021.
 PR 27-MAY-1999; 99US-0136392.
 PR 28-MAY-1999; 99US-0136782.
 PR 01-JUN-1999; 99US-0137222.
 PR 03-JUN-1999; 99US-0137528.
 PR 04-JUN-1999; 99US-0137502.
 PR 07-JUN-1999; 99US-0137724.
 PR 08-JUN-1999; 99US-0138094.
 PR 10-JUN-1999; 99US-0138540.
 PR 10-JUN-1999; 99US-0138847.
 PR 14-JUN-1999; 99US-0139119.
 PR 16-JUN-1999; 99US-0139452.
 PR 16-JUN-1999; 99US-0139453.
 PR 17-JUN-1999; 99US-0139452.
 PR 18-JUN-1999; 99US-0139454.
 PR 18-JUN-1999; 99US-0139455.
 PR 18-JUN-1999; 99US-0139456.
 PR 18-JUN-1999; 99US-0139457.
 PR 18-JUN-1999; 99US-0139458.
 PR 18-JUN-1999; 99US-0139459.
 PR 18-JUN-1999; 99US-0139460.
 PR 18-JUN-1999; 99US-0139461.
 PR 18-JUN-1999; 99US-0139462.
 PR 18-JUN-1999; 99US-0139463.
 PR 18-JUN-1999; 99US-0139750.
 PR 18-JUN-1999; 99US-0139763.
 PR 21-JUN-1999; 99US-0139817.
 PR 22-JUN-1999; 99US-0139899.
 PR 23-JUN-1999; 99US-0140353.
 PR 23-JUN-1999; 99US-0140354.
 PR 24-JUN-1999; 99US-0140695.
 PR 28-JUN-1999; 99US-0140823.
 PR 29-JUN-1999; 99US-0140991.
 PR 30-JUN-1999; 99US-0141287.
 PR 01-JUL-1999; 99US-0141842.
 PR 01-JUL-1999; 99US-0142154.
 PR 02-JUL-1999; 99US-0142055.
 PR 06-JUL-1999; 99US-0142390.
 PR 08-JUL-1999; 99US-0142803.
 PR 09-JUL-1999; 99US-0142920.
 PR 12-JUL-1999; 99US-0142977.
 PR 13-JUL-1999; 99US-0143342.
 PR 14-JUL-1999; 99US-0143624.

PR 15-JUL-1999; 99US-0144005.
 PR 16-JUL-1999; 99US-0144085.
 PR 16-JUL-1999; 99US-0144086.
 PR 19-JUL-1999; 99US-0144325.
 PR 19-JUL-1999; 99US-0144331.
 PR 19-JUL-1999; 99US-0144332.
 PR 19-JUL-1999; 99US-0144333.
 PR 19-JUL-1999; 99US-0144334.
 PR 19-JUL-1999; 99US-0144335.
 PR 20-JUL-1999; 99US-0144352.
 PR 20-JUL-1999; 99US-0144632.
 PR 20-JUL-1999; 99US-0144884.
 PR 21-JUL-1999; 99US-0144814.
 PR 21-JUL-1999; 99US-0145086.
 PR 21-JUL-1999; 99US-0145088.
 PR 22-JUL-1999; 99US-0145085.
 PR 22-JUL-1999; 99US-0145087.
 PR 22-JUL-1999; 99US-0145089.
 PR 23-JUL-1999; 99US-0145192.
 PR 23-JUL-1999; 99US-0145145.
 PR 23-JUL-1999; 99US-0145218.
 PR 23-JUL-1999; 99US-0145224.
 PR 26-JUL-1999; 99US-0145276.
 PR 27-JUL-1999; 99US-0145913.
 PR 27-JUL-1999; 99US-0145918.
 PR 27-JUL-1999; 99US-0145919.
 PR 28-JUL-1999; 99US-0145951.
 PR 02-AUG-1999; 99US-0146386.
 PR 02-AUG-1999; 99US-0146388.
 PR 02-AUG-1999; 99US-0146389.
 PR 03-AUG-1999; 99US-0147038.
 PR 04-AUG-1999; 99US-0147204.
 PR 04-AUG-1999; 99US-0147302.
 PR 05-AUG-1999; 99US-0147192.
 PR 05-AUG-1999; 99US-0147260.
 PR 06-AUG-1999; 99US-0147303.
 PR 06-AUG-1999; 99US-0147416.
 PR 09-AUG-1999; 99US-0147493.
 PR 09-AUG-1999; 99US-0147935.
 PR 10-AUG-1999; 99US-0148171.
 PR 11-AUG-1999; 99US-0148319.
 PR 12-AUG-1999; 99US-0148341.
 PR 13-AUG-1999; 99US-0148565.
 PR 13-AUG-1999; 99US-0148684.
 PR 16-AUG-1999; 99US-0149368.
 PR 17-AUG-1999; 99US-0149175.
 PR 18-AUG-1999; 99US-0149426.
 PR 20-AUG-1999; 99US-0149722.
 PR 20-AUG-1999; 99US-0149723.
 PR 20-AUG-1999; 99US-0149929.
 PR 23-AUG-1999; 99US-0149902.
 PR 23-AUG-1999; 99US-0149930.
 PR 25-AUG-1999; 99US-0150566.
 PR 26-AUG-1999; 99US-0150884.
 PR 27-AUG-1999; 99US-0151065.
 PR 27-AUG-1999; 99US-0151066.
 PR 30-AUG-1999; 99US-0151080.
 PR 31-AUG-1999; 99US-0151303.
 PR 31-AUG-1999; 99US-0151438.
 PR 01-SEP-1999; 99US-0151930.
 PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.

PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match 83.5%; Score 14.2; DB 21; Length 1427;
Best Local Similarity 76.5%; Pred. No. 76;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNA 17
||:|||||
Db 1161 gatttcggttggttaa 1177

RESULT 7
AAC33965
ID AAC33965 standard; DNA; 1429 bp.
XX AC AAC33965;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 4961.
XX KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-0301439.
XX PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137526.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138647.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139453.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139753.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.

```
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147935.
PR 09-AUG-1999; 99US-0147937.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 28-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
```

```
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

Query Match      83.5%; Score 14.2; DB 21; Length 1429;
Best Local Similarity 76.5%; Pred. No. 76;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 GAYTTYGGTGGGNA 17
      ||:||:||||| ||
Db      1163 gattcggtggggttaa 1179

RESULT      8
AAC64784
ID      AAC64784 standard; cDNA; 1436 BP.
XX
AC      AAC64784;
XX
DT      28-FEB-2001 (first entry)
XX
DE      Lemon alcohol acyl transferase encoding cDNA SEQ ID NO:20A.
XX
KW      Strawberry; fruit flavour; biosynthetic pathway; aliphatic; thiolase;
KW      aromatic ester; alcohol acyl transferase; alcohol dehydrogenase;
KW      pyruvate decarboxylase; aminotransferase; esterase; alcohol; aldehyde;
KW      alpha-keto acid; amino acid; fatty acid; acyl-CoA; processed food;
KW      food additive; flavour; syrup; ice-cream; frozen dessert; yoghurt;
KW      confectionery; flavouring; oral medication; vitamin; aroma; beverage;
KW      alcohol; scent; fragrance; perfume; cosmetic; suspension aid;
KW      aluminium salt; anti-perspirant; pharmaceutical; cleaning product;
KW      insect pheromone; dye carrier; solvent; insect repellent; miticide;
KW      scabicide; plasticiser; deodorant; ss.
XX
OS      Citrus limon.
XX
PN      WO200032789-Al.
XX
PD      08-JUN-2000.
XX
PF      02-DEC-1999; 99WO-NL00737.
XX
PR      02-DEC-1998; 98EP-0204018.
PR      12-MAR-1999; 99EP-0200739.
XX
PA      (CPRO-) CPRO-DLO CENT PLANTENVEREDELINGS REPROD.
XX
PI      Aharoni A, Luecker J, Verhoeven HA, Van Tunen AJ, O'Connell AP;
XX      WPI; 2000-412335/35.
DR      P-PSDB; AAB36458.
```

```

XX A new DNA sequence encoding a polypeptide with alcohol acyl transferase
PT activity for producing and regulating aromatic and/or aliphatic ester
PT formation in microorganisms, plant cells or plants
XX
XX Example 5; Page 114-115; 163pp; English.
XX
XX The present invention describes nucleotide sequences with thiolase,
CC alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase activities, which are involved in the
CC biosynthetic pathway for aliphatic and/or aromatic ester production in
CC fruit. The nucleotide sequences can be inserted into the genome of a
CC fruit-producing plant to regulate aliphatic and/or aromatic ester
CC formation. Aromatic and/or aliphatic esters in microorganisms, plant
CC cells or plants are produced by inserting thiolase, alcohol acyl
CC transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase nucleotide sequences into the genome and
CC feeding the microorganism or plant with alcohol, aldehydes, alpha-keto
CC acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and
CC their proteins can be used in the processed food industry as food
CC additives to enhance the flavour of syrups, ice-creams, frozen desserts,
CC yoghurts and confectionery. They are used: as flavouring agents for oral
CC medications and vitamins; provide flavour and aroma in beverages,
CC including alcohol; enhance or reduce fruit flavour, aroma, fragrance or
CC scent; enhance the flavour or aroma of natural, synthetic or artificial
CC products; for the production of novel combinations of artificial flavour
CC substances; as antibacterial or anti-fungal agents; as fragrance or
CC perfumes in cosmetics, creams, sun-protectant products, hair
CC conditioners, lengthening agents and fixatives in perfumes, suspension
CC aids for aluminium salts in anti-perspirant pharmaceuticals, cleaning
CC products, personal care products and animal care products; as
CC disinfectant additives; as degreasing solvents for electronics; as
CC insect pheromones; and as dye carriers, solvents, insect repellents,
CC miticides, scabicides, plasticisers and deodorants. The present sequence
CC encodes lemon alcohol acyl transferase, from the present invention.
XX
XX Sequence 1436 BP; 410 A; 280 C; 320 G; 436 T; 0 other;

Query Match      83.5%; Score 14.2; DB 21; Length 1436;
Best Local Similarity 76.5%; Pred. No. 76;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGGNAA 17
Db 1126 gatttcgggtgggggaa 1142
    ||:|:|:|:|:|:|:|:|:|

RESULT 9
AAC64787
ID AAC64787 standard; cDNA; 1470 BP.
XX
AC AAC64787;
XX
XX 28-FEB-2001 (first entry)
XX
XX Honey dew melon alcohol acyl transferase encoding cDNA SEQ ID NO:23A.
XX
XX Strawberry; fruit flavour; biosynthetic pathway; aliphatic; thiolase;
KW aromatic ester; alcohol acyl transferase; alcohol dehydrogenase;
KW pyruvate decarboxylase; aminotransferase; esterase; alcohol; aldehyde;
KW alpha-keto acid; amino acid; fatty acid; acyl-CoA; processed food;
KW food additive; flavour; syrup; ice-cream; frozen dessert; yoghurt;
KW confectionery; flavouring; oral medication; vitamin; aroma; beverage;
KW alcohol; scent; fragrance; perfume; cosmetic; suspension aid;
KW aluminium salt; anti-perspirant; pharmaceutical; cleaning product;
KW insect pheromone; dye carrier; solvent; insect repellent; miticide;
KW scabicide; plasticiser; deodorant; ss.
XX
XX Cucumis sp.
OS
XX WO2000032789-A1.
PN
XX

```

```

PD 08-JUN-2000.
XX
XX 02-DEC-1999; 99WO-NL00737.
XX
XX 02-DEC-1999; 98EP-0204018.
XX
XX 12-MAR-1999; 99EP-0200739.
XX
XX (CPRO-) CPRO-DLO CENT PLANTENVERDEDELINGS REPROD.
XX
XX Aharoni A, Luecker J, Verhoeven HA, Van Tunen AJ, O'Connell AP;
XX WPI; 2000-412335/35.
XX P-PSDB; AAB36461.
XX
XX A new DNA sequence encoding a polypeptide with alcohol acyl transferase
XX activity for producing and regulating aromatic and/or aliphatic ester
XX formation in microorganisms, plant cells or plants
XX
XX Example 5; Page 121-122; 163pp; English.
XX
XX The present invention describes nucleotide sequences with thiolase,
CC alcohol acyl transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase activities, which are involved in the
CC biosynthetic pathway for aliphatic and/or aromatic ester production in
CC fruit. The nucleotide sequences can be inserted into the genome of a
CC fruit-producing plant to regulate aliphatic and/or aromatic ester
CC formation. Aromatic and/or aliphatic esters in microorganisms, plant
CC cells or plants are produced by inserting thiolase, alcohol acyl
CC transferase, alcohol dehydrogenase, pyruvate decarboxylase,
CC aminotransferase and esterase nucleotide sequences into the genome and
CC feeding the microorganism or plant with alcohol, aldehydes, alpha-keto
CC acids, or amino acids and fatty acids, and acyl-CoA. The nucleotides and
CC their proteins can be used in the processed food industry as food
CC additives to enhance the flavour of syrups, ice-creams, frozen desserts,
CC yoghurts and confectionery. They are used: as flavouring agents for oral
CC medications and vitamins; provide flavour and aroma in beverages,
CC including alcohol; enhance or reduce fruit flavour, aroma, fragrance or
CC scent; enhance the flavour or aroma of natural, synthetic or artificial
CC products; for the production of novel combinations of artificial flavour
CC substances; as antibacterial or anti-fungal agents; as fragrance or
CC perfumes in cosmetics, creams, sun-protectant products, hair
CC conditioners, lengthening agents and fixatives in perfumes, suspension
CC aids for aluminium salts in anti-perspirant pharmaceuticals, cleaning
CC products, personal care products and animal care products; as
CC disinfectant additives; as degreasing solvents for electronics; as
CC insect pheromones; and as dye carriers, solvents, insect repellents,
CC miticides, scabicides, plasticisers and deodorants. The present sequence
CC encodes honey dew melon alcohol acyl transferase, from the present
XX invention.
XX
XX Sequence 1470 BP; 415 A; 301 C; 334 G; 420 T; 0 other;

Query Match      83.5%; Score 14.2; DB 21; Length 1470;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGGNAA 17
Db 1122 gactttggatggggaaa 1138
    ||:|:|:|:|:|:|:|:|:|

RESULT 10
AAC39128
ID AAC39128 standard; DNA; 1477 BP.
XX
AC AAC39128;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 23476.
DE
XX Hybridisation assay; genetic mapping; gene expression control;
XX

```

KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.

KW	protein identification; signal transduction pathway;	PR	24-JUN-1999;	99US-0140695;
KW	metabolic pathway; promoter; termination sequence; ss.	PR	28-JUN-1999;	99US-0140823;
XX		PR	29-JUN-1999;	99US-0140991;
OS		PR	30-JUN-1999;	99US-0141287;
XX	Arabidopsis thaliana.	PR	01-JUL-1999;	99US-0141842;
PN		PR	01-JUL-1999;	99US-0142154;
XX	EP1033405-A2.	PR	01-JUL-1999;	99US-0142055;
PD		PR	02-JUL-1999;	99US-0142390;
XX	06-SEP-2000.	PR	06-JUL-1999;	99US-0142803;
XX		PR	08-JUL-1999;	99US-0142920;
PF		PR	09-JUL-1999;	99US-0142977;
XX	25-FEB-2000; 2000EP-0301439.	PR	12-JUL-1999;	99US-0143542;
PR		PR	13-JUL-1999;	99US-0143542;
XX	25-FEB-1999;	PR	14-JUL-1999;	99US-0143624;
PR	05-MAR-1999;	PR	15-JUL-1999;	99US-0144005;
PR	09-MAR-1999;	PR	16-JUL-1999;	99US-0144085;
PR	23-MAR-1999;	PR	16-JUL-1999;	99US-0144086;
PR	25-MAR-1999;	PR	19-JUL-1999;	99US-0144325;
PR	29-MAR-1999;	PR	19-JUL-1999;	99US-0144331;
PR	01-APR-1999;	PR	19-JUL-1999;	99US-0144332;
PR	06-APR-1999;	PR	19-JUL-1999;	99US-0144333;
PR	08-APR-1999;	PR	19-JUL-1999;	99US-0144334;
PR	16-APR-1999;	PR	19-JUL-1999;	99US-0144335;
PR	19-APR-1999;	PR	20-JUL-1999;	99US-0144352;
PR	21-APR-1999;	PR	20-JUL-1999;	99US-0144632;
PR	23-APR-1999;	PR	20-JUL-1999;	99US-0144884;
PR	28-APR-1999;	PR	21-JUL-1999;	99US-0144814;
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145086;
PR	30-APR-1999;	PR	21-JUL-1999;	99US-0145088;
PR	04-MAY-1999;	PR	21-JUL-1999;	99US-0145085;
PR	05-MAY-1999;	PR	22-JUL-1999;	99US-0145087;
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145089;
PR	06-MAY-1999;	PR	22-JUL-1999;	99US-0145192;
PR	07-MAY-1999;	PR	23-JUL-1999;	99US-0145145;
PR	11-MAY-1999;	PR	23-JUL-1999;	99US-0145218;
PR	14-MAY-1999;	PR	23-JUL-1999;	99US-0145224;
PR	14-MAY-1999;	PR	26-JUL-1999;	99US-0145276;
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145913;
PR	14-MAY-1999;	PR	27-JUL-1999;	99US-0145918;
PR	18-MAY-1999;	PR	27-JUL-1999;	99US-0145919;
PR	19-MAY-1999;	PR	28-JUL-1999;	99US-0145951;
PR	20-MAY-1999;	PR	02-AUG-1999;	99US-0146386;
PR	21-MAY-1999;	PR	02-AUG-1999;	99US-0146388;
PR	24-MAY-1999;	PR	02-AUG-1999;	99US-0146389;
PR	25-MAY-1999;	PR	03-AUG-1999;	99US-0147038;
PR	27-MAY-1999;	PR	04-AUG-1999;	99US-0147204;
PR	28-MAY-1999;	PR	04-AUG-1999;	99US-0147302;
PR	01-JUN-1999;	PR	05-AUG-1999;	99US-0147192;
PR	03-JUN-1999;	PR	05-AUG-1999;	99US-0147260;
PR	04-JUN-1999;	PR	06-AUG-1999;	99US-0147303;
PR	07-JUN-1999;	PR	06-AUG-1999;	99US-0147416;
PR	08-JUN-1999;	PR	09-AUG-1999;	99US-0147493;
PR	10-JUN-1999;	PR	09-AUG-1999;	99US-0147935;
PR	10-JUN-1999;	PR	10-AUG-1999;	99US-0148171;
PR	14-JUN-1999;	PR	11-AUG-1999;	99US-0148319;
PR	16-JUN-1999;	PR	12-AUG-1999;	99US-0148341;
PR	16-JUN-1999;	PR	13-AUG-1999;	99US-0148565;
PR	17-JUN-1999;	PR	13-AUG-1999;	99US-0148684;
PR	18-JUN-1999;	PR	16-AUG-1999;	99US-0149368;
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175;
PR	18-JUN-1999;	PR	17-AUG-1999;	99US-0149175;
PR	18-JUN-1999;	PR	18-AUG-1999;	99US-0149426;
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149722;
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149723;
PR	18-JUN-1999;	PR	20-AUG-1999;	99US-0149929;
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149902;
PR	18-JUN-1999;	PR	23-AUG-1999;	99US-0149930;
PR	18-JUN-1999;	PR	25-AUG-1999;	99US-0150566;
PR	18-JUN-1999;	PR	26-AUG-1999;	99US-0150884;

PR 07-SEP-1999; 99US-0152363.
 PR 10-SEP-1999; 99US-0153070.
 PR 13-SEP-1999; 99US-0153758.
 PR 15-SEP-1999; 99US-0154018.
 PR 16-SEP-1999; 99US-0154039.
 PR 20-SEP-1999; 99US-0154779.
 PR 22-SEP-1999; 99US-0155139.
 PR 23-SEP-1999; 99US-0155486.
 PR 24-SEP-1999; 99US-0155659.
 PR 28-SEP-1999; 99US-0156458.
 PR 29-SEP-1999; 99US-0156596.
 PR 04-OCT-1999; 99US-0157117.
 PR 05-OCT-1999; 99US-0157753.
 PR 06-OCT-1999; 99US-0157865.
 PR 07-OCT-1999; 99US-0158029.
 PR 08-OCT-1999; 99US-0158232.
 PR 12-OCT-1999; 99US-0158369.
 PR 13-OCT-1999; 99US-0159293.
 PR 13-OCT-1999; 99US-0159294.
 PR 13-OCT-1999; 99US-0159295.
 PR 14-OCT-1999; 99US-0159329.
 PR 14-OCT-1999; 99US-0159330.
 PR 14-OCT-1999; 99US-0159331.
 PR 14-OCT-1999; 99US-0159637.
 PR 14-OCT-1999; 99US-0159638.
 PR 18-OCT-1999; 99US-0159584.
 PR 21-OCT-1999; 99US-0160741.
 PR 21-OCT-1999; 99US-0160767.
 PR 21-OCT-1999; 99US-0160768.
 PR 21-OCT-1999; 99US-0160770.
 PR 21-OCT-1999; 99US-0160814.
 PR 21-OCT-1999; 99US-0160815.
 PR 22-OCT-1999; 99US-0160980.
 PR 22-OCT-1999; 99US-0160981.
 PR 22-OCT-1999; 99US-0160989.
 PR 25-OCT-1999; 99US-0161404.
 PR 25-OCT-1999; 99US-0161405.
 PR 25-OCT-1999; 99US-0161406.
 PR 26-OCT-1999; 99US-0161359.
 PR 26-OCT-1999; 99US-0161360.
 PR 26-OCT-1999; 99US-0161361.
 PR 28-OCT-1999; 99US-0161920.
 PR 28-OCT-1999; 99US-0161992.
 PR 28-OCT-1999; 99US-0161993.
 PR 29-OCT-1999; 99US-0162142.

Query Match 83.5%; Score 14.2; DB 21; Length 1477;
 Best Local Similarity 76.5%; Pred. No. 77;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
 ||:||||| ||||| ||
 Db 1186 gatttcggttggtggttaa 1202

RESULT 11
 AAT37311
 ID AAT37311 standard; cDNA to mRNA; 1479 BP.
 XX AC AAT37311;
 XX DT 06-FEB-1997 (first entry)
 XX DE Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
 KW Senecio cruentus; Lavandula angustifolia; ds.
 XX OS Perilla ocimoides (Clone pSAT208).
 XX PN
 XX PD Key Location/Qualifiers

FT CDS 3..1343
 FT /*tag= a
 FT /product= Aromatic acyl transferase.

PN WO9625500-A1.

PD 22-AUG-1996.

PF 16-FEB-1996; 96WO-JP00348.

PR 30-JAN-1996; 96JP-0046534.

PR 17-FEB-1995; 95JP-0067159.

PR 29-JUN-1995; 95JP-0196915.

XX (SUNR) SUNTORY LTD.

XX Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;

PI Nakao M, Tanaka Y, Yonekura K;

XX WPI: 1996-393401/39.

DR P-PSDB; AA04725.

XX DNA coding for aromatic acyl transferase - for transforming plants
 PT which produce anthocyanin pigments and thus altering colour tone,
 PT e.g. of flowers

PS Claim 4; Page 65-69; 94pp; Japanese.

XX Vectors containing DNA fragments encoding proteins of plant origin
 CC with aromatic acyl transferase activity may be used to transform
 CC plants which produce anthocyanin pigments. The aromatic acyl
 CC transferase acylates the pigments in the flower resulting in colour
 CC tone changes and allowing new colourations to be produced. Six
 CC specific DNA sequences encoding aromatic acyl transferase from
 CC different plants are described in AAT37308-T37313.

XX Sequence 1479 BP; 420 A; 316 C; 331 G; 411 T; 1 other;

Query Match 83.5%; Score 14.2; DB 17; Length 1479;
 Best Local Similarity 76.5%; Pred. No. 77;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
 ||:||||| ||||| ||

Db 1167 gatttcggttggtggttaa 1183

RESULT 12
 AAT37312
 ID AAT37312 standard; cDNA to mRNA; 1508 BP.
 XX AC AAT37312;
 XX DT 06-FEB-1997 (first entry);
 XX DE Aromatic acyl transferase coding sequence.

XX Aromatic acyl transferase; transformation; anthocyanin pigment;
 KW plants; acylation; colour; tone; colouration; colour change;
 KW Gentiana triflora; Petunia hybrida; Perilla ocimoides;
 KW Senecio cruentus; Lavandula angustifolia; ds.
 XX OS Senecio cruentus (Clone pCAT8).
 XX PN
 XX PD Key Location/Qualifiers

FT CDS 3..1367

FT /*tag= a

FT /product= Aromatic acyl transferase.

XX WO9625500-A1.

XX 22-AUG-1996.

```
XX PF 16-FEB-1996; 96WO-JP00348.
XX PR 30-JAN-1996; 96JP-0046534.
XX PR 17-FEB-1995; 95JP-0067159.
XX PR 29-JUN-1995; 95JP-0196915.
XX PA (SUNR ) SUNTORY LTD.
XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
XX PI Nakao M, Tanaka Y, Yonekura K;
XX DR WPI; 1996-393401/39.
XX PS P-PSDB; AAW04726.
XX PT DNA coding for aromatic acyl transferase - for transforming plants
XX PT which produce anthocyanin pigments and thus altering colour tone,
XX PT e.g. of flowers
XX PS Claim 4; Page 69-72; 94pp; Japanese.
XX CC Vectors containing DNA fragments encoding proteins of plant origin
XX CC with aromatic acyl transferase activity may be used to transform
XX CC plants which produce anthocyanin pigments. The aromatic acyl
XX CC transferase acylates the pigments in the flower resulting in colour
XX CC tone changes and allowing new colourations to be produced. Six
XX CC specific DNA sequences encoding aromatic acyl transferase from
XX CC different plants are described in AAT37308-T37313.
XX SQ Sequence 1508 BP; 442 A; 294 C; 295 G; 477 T; 0 other;

Query Match 83.5%; Score 14.2; DB 17; Length 1508;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNAA 17
Db 1182 gatttgggtgggga 1198
||:|:|:|:|:|:|

RESULT 13
AAT37313
ID AAT37313 standard; cDNA to mRNA; 1518 BP.
AC AAT37313;
XX DT 06-FEB-1997 (first entry)
XX DE Aromatic acyl transferase coding sequence.
XX KW Aromatic acyl transferase; transformation; anthocyanin pigment;
XX KW plants; acylation; colour; tone; colouration; colour change;
XX KW Gentiana triflora; Petunia hybrida; Perilla occimoides;
XX KW Scenecio cruentus; Lavandula angustifolia; ds.
XX OS Lavandula angustifolia (Clone pLAT21).
XX PH Key Location/Qualifiers
XX FT CDS 1..1355
XX FT /*tag= a
XX FT /product= Aromatic acyl transferase.
XX PN WO9625500-A1.
XX PD 22-AUG-1996.
XX PF 16-FEB-1996; 96WO-JP00348.
XX PR 30-JAN-1996; 96JP-0046534.
XX PR 17-FEB-1995; 95JP-0067159.
XX PR 29-JUN-1995; 95JP-0196915.
XX
```

```
PA (SUNR ) SUNTORY LTD.
XX PI Ashikari T, Fujiwara H, Fukui Y, Kusumi T, Mizutani M;
XX PI Nakao M, Tanaka Y, Yonekura K;
XX DR WPI; 1996-393401/39.
XX PS DNA coding for aromatic acyl transferase - for transforming plants
XX PT which produce anthocyanin pigments and thus altering colour tone,
XX PT e.g. of flowers
XX PS Claim 4; Page 73-76; 94pp; Japanese.
XX CC Vectors containing DNA fragments encoding proteins of plant origin
XX CC with aromatic acyl transferase activity may be used to transform
XX CC plants which produce anthocyanin pigments. The aromatic acyl
XX CC transferase acylates the pigments in the flower resulting in colour
XX CC tone changes and allowing new colourations to be produced. Six
XX CC specific DNA sequences encoding aromatic acyl transferase from
XX CC different plants are described in AAT37308-T37313. NOTE: This
XX CC sequence is supposed to cross reference with the protein described
XX CC in AAW04727, however there are so many discrepancies between the
XX CC polypeptide decoded from this sequence and the polypeptide given in
XX CC the specification and described in AAW04727 that the indexer decided
XX CC not to cross reference the two.
XX SQ Sequence 1518 BP; 384 A; 340 C; 366 G; 428 T; 0 other;

Query Match 83.5%; Score 14.2; DB 17; Length 1518;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGTGGGNAA 17
Db 1176 gatttgggtgggtaa 1192
||:|:|:|:|:|:|

RESULT 14
AAT89415
ID AAT89415 standard; cDNA; 1526 BP.
XX AC AAT89415;
XX DT 15-APR-1998 (first entry)
XX DE Melon ripening-related cDNA clone MEL2.
XX KW Melon cultivar Canteloupe charentais; fruit ripening control;
XX KW ethylene; MEL2 promoter; ss.
XX OS Cucumis melo.
XX PH Key Location/Qualifiers
XX FT 3'UTR 1385..1526
XX FT /*tag= a
XX FT /note= "contains a putative polyadenylation signal"
XX PN WO9737023-A1.
XX PD 09-OCT-1997.
XX PF 24-MAR-1997; 97WO-GB00824.
XX PR 02-APR-1996; 96GB-0006906.
XX PA (ZENE ) ZENECA LTD.
XX PI Aggelis A, Greierson D, John I, Karvouni Z;
XX DR WPI; 1997-503108/46.
XX PT cDNA related with fruit ripening - can be used to genetically modify
```

```

PT especially melons to control ripening
PS Claim 1: Pages 18-19; 38pp; English.
XX
XX This cDNA sequence is from the MEL2 clone which produces a
CC novel ripening-related product from Cucumis melo. It is one of two
CC specified sequences (MEL2) 1526 bp or (MEL7) 686 bp which were
CC isolated from a melon ripe fruit cDNA library. MEL2 is not a
CC full-length clone; it lacks the initiation codon for the amino terminus.
CC The cDNA allows control of the ripening of fruit, especially melons.
XX
XX Sequence 1526 BP; 439 A; 297 C; 331 G; 448 T; 2 U; 9 other;
SQ

```

```

Query Match      83.5%; Score 14.2; DB 15; Length 1542;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
   ||:||:| | |||| | |
Db 1136 gactttggtgaggaaa 1152

RESULT 15
AAQ54685
ID AAQ54685 standard; cDNA; 1542 BP.
XX
AC AAQ54685;
XX
DT 06-JUL-1994 (first entry)
XX
DE Anther specific cDNA clone ant32.
XX
KW Transgenic plants; male sterility; pollen; sterile;
KW self-pollination; seed; hybrid; toxin-A; peIE; T-urf13; Gin; iaal;
KW CytA; toxin; Nicotiana tabacum; Diptheria; Erwinia chrysanthemi;
KW phage Mu; Psuedomonas syringae; Bacillus thuringiensis; anther;
KW breeding; ss.
XX
OS Nicotiana tabacum.
XX
FH Key Location/Qualifiers
FT CDS 66..1412
FT FT /*tag= a
XX
XX EP578611-A.
XX
PD 12-JAN-1994.
XX
XX 24-JUN-1993; 93EP-0810455.
XX
PR 02-JUL-1992; 92US-0908242.
XX
XX (CIBA ) CIBA GEIGY AG.
XX
XX Crossland LD, Tuttle AB;
XX
XX WPI: 1994-010428/02.
XX
XX P-PSDB; AAR47475.
XX
XX Anther-specific cDNA, genomic and recombinant DNA - produce
PT transgenic male-sterile plants, which prevents self-pollination,
PT in hybrid seed prodn.
XX
XX Claim 1: Page 21-24; 75pp; English.
XX
XX Anther specific cDNA or genomic sequences can be used to identify
CC and isolate anther specific promoters. The anther specific promoter
CC can then be cloned into a recombinant construct and used to express
CC heterologous genes. Preferred heterologous genes include Diptheria
CC toxin A-chain gene; pectate lyase gene peIE from Erwinia
CC chrysanthemi; T-urf13 from cms-T maize mitochondrial genomes; the
CC Gin recombinase gene from phage Mu; the indole acetic acid-lysine

```

```

CC synthetase gene from Pseudomonas syringae and the CytA toxin gene
CC from Bacillus thuringiensis israeliensis. All of these genes when
CC expressed in anther tissue will result in the inability of the
CC plant to produce viable pollen. Transformation of plants with such
CC a recombinant construct can produce transgenic, male sterile plants.
CC Male sterility is important in the production of hybrid seeds as it
CC prevents self pollination which hinders breeding and hybrid seed
CC production.
XX
XX Sequence 1542 BP; 473 A; 293 C; 332 G; 444 T; 0 other;
SQ

```

```

Query Match      83.5%; Score 14.2; DB 15; Length 1542;
Best Local Similarity 76.5%; Pred. No. 77;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNAA 17
   ||:||:| | |||| | |
Db 1239 gattttgggtgaggaaa 1255

Search completed: November 5, 2001, 18:12:09
Job time: 15464 sec

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:05:10 ; Search time 168.74 Seconds
(without alignments)
19.072 Million cell updates/sec

Title: US-08-894-356C-22
Perfect score: 17
Sequence: 1 GAYTTYGGNTGGGNA 17

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 324599 seqs, 9465562 residues

Total number of hits satisfying chosen parameters: 649198

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	14.2	83.5	23	US-09-187-050-6	Sequence 6, Appli
2	14.2	83.5	1080	US-07-598-873-1	Sequence 1, Appli
3	14.2	83.5	1080	US-08-073-425-1	Sequence 1, Appli
4	14.2	83.5	1080	US-08-396-531-1	Sequence 1, Appli
5	14.2	83.5	1096	US-09-056-153-1	Sequence 1, Appli
6	14.2	83.5	1512	US-09-142-514-1	Sequence 1, Appli
7	14.2	83.5	1542	US-08-207-904-1	Sequence 1, Appli
8	14.2	83.5	3706	US-08-207-904-16	Sequence 16, Appli
9	13	76.5	4089	US-07-908-245-1	Sequence 1, Appli
C 10	13	76.5	4097	US-09-123-708-5	Sequence 5, Appli
C 11	13	76.5	4097	US-09-123-624-5	Sequence 5, Appli
12	12.6	74.1	508	US-08-822-028-74	Sequence 74, Appli
13	12.6	74.1	508	US-08-478-285-74	Sequence 74, Appli
14	12.6	74.1	912	US-08-593-171A-24	Sequence 24, Appli
15	12.6	74.1	912	US-08-646-590B-24	Sequence 24, Appli
16	12.6	74.1	912	US-09-069-226-24	Sequence 24, Appli
17	12.6	74.1	912	US-09-413-184-24	Sequence 24, Appli
18	12.6	74.1	930	US-08-096-182A-1	Sequence 1, Appli
19	12.6	74.1	930	US-08-877-109-1	Sequence 1, Appli
20	12.6	74.1	930	US-08-798-760-1	Sequence 1, Appli
21	12.6	74.1	930	PCT-US94-08327-1	Sequence 1, Appli
22	12.6	74.1	2344	US-07-695-472B-35	Sequence 35, Appli
23	12.6	74.1	2347	US-08-250-740-32	Sequence 32, Appli
24	12.6	74.1	2347	US-07-695-472B-1	Sequence 1, Appli
25	12.6	74.1	2347	US-07-693-472B-34	Sequence 34, Appli
26	12.6	74.1	2347	US-07-695-472B-36	Sequence 36, Appli
27	12.6	74.1	2560	US-07-916-098A-44	Sequence 44, Appli

ALIGNMENTS

RESULT 1

US-09-187-050-6/c
; Sequence 6, Application US/09187050B
; Patent No. 6043072
; GENERAL INFORMATION:
; APPLICANT: Croteau, Rodney B
; APPLICANT: Hefner, Jerry
; TITLE OF INVENTION: Nucleic Acids Encoding Taxus Geranylgeranyl Diphosphate
; FILE REFERENCE: WSU12423
; CURRENT APPLICATION NUMBER: US/09/187,050B
; CURRENT FILING DATE: 1998-11-05
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 23
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: non-degenerate
; OTHER INFORMATION: reverse PCR primer
; NAME/KEY: misc_difference
; LOCATION: (1)..(23)
; OTHER INFORMATION: No. 6043072-degenerate PCR primer

US-09-187-050-6
Query Match 83.5%; Score 14.2; DB 3; Length 23;
Best Local Similarity 76.5%; Pred. No. 8.4;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|||||
DB 20 GATTTGGCTGGGGTAA 4
||:|||||

RESULT 2
US-07-598-873-1
; Sequence 1, Application US/07598873
; Patent No. 5254800
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

US-09-187-050-6
Query Match 83.5%; Score 14.2; DB 3; Length 23;
Best Local Similarity 76.5%; Pred. No. 8.4;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|||||
DB 20 GATTTGGCTGGGGTAA 4
||:|||||

RESULT 2
US-07-598-873-1
; Sequence 1, Application US/07598873
; Patent No. 5254800
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

US-09-187-050-6
Query Match 83.5%; Score 14.2; DB 3; Length 23;
Best Local Similarity 76.5%; Pred. No. 8.4;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|||||
DB 20 GATTTGGCTGGGGTAA 4
||:|||||

RESULT 2
US-07-598-873-1
; Sequence 1, Application US/07598873
; Patent No. 5254800
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS
; TITLE OF INVENTION: DERIVED THEREFROM
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:

US-09-187-050-6
Query Match 83.5%; Score 14.2; DB 3; Length 23;
Best Local Similarity 76.5%; Pred. No. 8.4;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|||||
DB 20 GATTTGGCTGGGGTAA 4
||:|||||

```
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA: US/07/598,873
; APPLICATION NUMBER: 07/598,873
; FILING DATE: 19901019
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 1080 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum
; STRAIN: Ailisa Craig
; DEVELOPMENTAL STAGE: Ripening
; US-07-598-873-1

Query Match      83.5%; Score 14.2; DB 1; Length 1080;
Best Local Similarity 76.5%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GAVTTYGGTGGGNA 17
   ||:|||||
Db 866 GAVTTGGATGGGAAA 882

RESULT 3
US-08-073-425-1
; Sequence 1, Application US/08073425
; Patent No. 5569829
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: BONIWELL, JEREMY M.
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: 1100 New York Avenue, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/425,425
; FILING DATE: 09-JUN-1993
; CLASSIFICATION: 800

; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Eleventh Floor, 1615 L Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20036-5601
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/598,873
; FILING DATE: 19901019
; CLASSIFICATION: 800
; ATTORNEY/AGENT INFORMATION:
; NAME: WILSON, MARY J
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 1080 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna to mRNA
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Lycopersicon esculentum
; STRAIN: Ailisa Craig
; DEVELOPMENTAL STAGE: Ripening
; US-07-598-873-1

Query Match      83.5%; Score 14.2; DB 1; Length 1080;
Best Local Similarity 76.5%; Pred. No. 14;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 GAVTTYGGTGGGNA 17
   ||:|||||
Db 866 GAVTTGGATGGGAAA 882

RESULT 4
US-08-396-531-1
; Sequence 1, Application US/08396531
; Patent No. 5744364
; GENERAL INFORMATION:
; APPLICANT: BIRD, COLIN R
; APPLICANT: GRIERSON, DONALD
; APPLICANT: RAY, JOHN A
; APPLICANT: SCHUCH, WOLFGANG W
; TITLE OF INVENTION: DNA, CONSTRUCTS, CELLS AND PLANTS
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CUSHMAN, DARBY & CUSHMAN
; STREET: Ninth Floor, 1100 New York Avenue, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/396,531
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/847,037
; FILING DATE: 16-APR-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: KOKULIS, PAUL N.
; REGISTRATION NUMBER: 16,773
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3000
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUSH
; INFORMATION FOR SEQ ID NO: 1:
; LENGTH: 1080 base pairs
; TYPE: nucleic acid
```

STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: CDNA to mRNA
 HYPOTHETICAL: NO
 ORIGINAL SOURCE:
 ORGANISM: Lycopersicon esculentum
 STRAIN: Ailisa Craig
 DEVELOPMENTAL STAGE: Ripening
 US-08-396-531-1

Query Match 83.5%; Score 14.2; DB 1; Length 1080;
 Best Local Similarity 76.5%; Pred. No. 14;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
 ||:||||| ||||| ||

Db 866 GAYTTGGATGGGAAA 882

RESULT 5

US-09-056-153-1
 ; Sequence 1, Application US/09056153
 ; Patent No. 6127113
 ; GENERAL INFORMATION:
 ; APPLICANT: Atkinson, Richard L.
 ; APPLICANT: Dhurandhar, Nikhil V.
 ; TITLE OF INVENTION: Viral Obesity Methods and Compositions
 ; FILE REFERENCE: 710395-90010
 ; CURRENT APPLICATION NUMBER: US/09/056,153
 ; CURRENT FILING DATE: 1998-04-06
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1096
 ; TYPE: DNA
 ; ORGANISM: Adenovirus type 36P
 US-09-056-153-1

Query Match 83.5%; Score 14.2; DB 3; Length 1096;
 Best Local Similarity 76.5%; Pred. No. 14;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
 ||:||||| ||||| ||

Db 1013 gactttggatggggttaa 1029

RESULT 6

US-09-142-514-1
 ; Sequence 1, Application US/09142514A
 ; Patent No. 6107548
 ; GENERAL INFORMATION:
 ; APPLICANT: Aggelis, Alexandros
 ; APPLICANT: Grierson, Donald
 ; APPLICANT: John, Isaac
 ; APPLICANT: Karvouni, Zoi
 ; TITLE OF INVENTION: Fruit Ripening
 ; FILE REFERENCE: SEES0111/UST
 ; CURRENT APPLICATION NUMBER: US/09/142,514A
 ; CURRENT FILING DATE: 1998-09-09
 ; EARLIER APPLICATION NUMBER: PCT/GB97/00824
 ; EARLIER FILING DATE: 1997-03-24
 ; EARLIER APPLICATION NUMBER: GB 9606906.7
 ; EARLIER FILING DATE: 1996-04-02
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 1512
 ; TYPE: DNA
 ; ORGANISM: Cucumis melo
 US-09-142-514-1

Query Match 83.5%; Score 14.2; DB 3; Length 1512;
 Best Local Similarity 76.5%; Pred. No. 15;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
 ||:||||| ||||| ||

Db 1122 gactttggatggggttaa 1138

RESULT 7

US-08-207-904-1
 ; Sequence 1, Application US/08207904
 ; Patent No. 5477002
 ; GENERAL INFORMATION:
 ; APPLICANT: Tuttle, AnnMarie
 ; APPLICANT: Crossland, Lyle D.
 ; TITLE OF INVENTION: Anther-Specific cDNA Sequences, Genomic
 ; TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
 ; NUMBER OF SEQUENCES: 21
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CIBA-GEIGY Corporation
 ; STREET: 7 Skyline Drive
 ; CITY: Hawthorne
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10532
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/207,904
 ; FILING DATE:
 ; CLASSIFICATION: 800
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: US/07/908,242
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Lazar, Steven R.
 ; REGISTRATION NUMBER: 32,618
 ; REFERENCE/DOCKET NUMBER: CGC 1624
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (919)541-8615
 ; TELEFAX: (919)541-8689
 ; INFORMATION FOR SEQ ID NO: 1:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 1542 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: CDNA
 ; HYPOTHETICAL: NO
 ; ANTI-SENSE: NO
 ; ORIGINAL SOURCE:
 ; ORGANISM: Nicotiana tabacum
 ; INDIVIDUAL ISOLATE: Ant32
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 66..1412
 US-08-207-904-1

Query Match 83.5%; Score 14.2; DB 1; Length 1542;
 Best Local Similarity 76.5%; Pred. No. 15;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
 ||:||||| ||||| ||

Db 1239 GAYTTGGTGGGAAA 1255

RESULT 8
US-08-207-904-16
; Sequence 16, Application US/08207904
; Patent No. 5477002
; GENERAL INFORMATION:
; APPLICANT: Tuttle, AnnMarie
; APPLICANT: Crossland, Lyle D.
; TITLE OF INVENTION: Anther-Specific cDNA Sequences, Genomic
; TITLE OF INVENTION: DNA Sequences and Recombinant DNA Sequences
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: CIBA-GEIGY Corporation
; STREET: 7 Skyline Drive
; CITY: Hawthorne
; STATE: New York
; COUNTRY: USA
; ZIP: 10532
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/207,904
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,242
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Lazar, Steven R.
; REGISTRATION NUMBER: 32,618
; REFERENCE/DOCKET NUMBER: CGC 1624
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919)541-8615
; TELEFAX: (919)541-8689
; INFORMATION FOR SEQ ID NO: 16:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3706 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Nicotiana tabacum
; INDIVIDUAL ISOLATE: Anc32 genomic clone
; IMMEDIATE SOURCE:
; CLONE: pCIB950
; FEATURE:
; NAME/KEY: TATA_signal
; LOCATION: 1971..1975
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 2076..3422
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 2009
; OTHER INFORMATION: /note= "Putative transcription
; OTHER INFORMATION: start site"
US-08-207-904-16

Query Match 83.5%; Score 14.2; DB 1; Length 3706;
Best Local Similarity 76.5%; Pred. No. 17;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNA 17
|||:|||||
Db 3249 GATTTGGTGGGAAA 3265

RESULT 9
US-07-908-245-1/c
; Sequence 1, Application US/07908245
; Patent No. 5498539
; GENERAL INFORMATION:
; APPLICANT: Harrison, David G.
; APPLICANT: Alexander, R. Wayne
; APPLICANT: Murphy, T.J.
; APPLICANT: Nishida, Ken'ichi
; TITLE OF INVENTION: Endothelial Nitric Oxide Synthase
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 1100 Peachtree Street, Suite 2800
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30309-4530
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/908,245
; FILING DATE: 19920702
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: EMU 111
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-815-6508
; TELEFAX: 404-815-6555
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 4089 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Bovine
; TISSUE TYPE: Aorta
; CELL TYPE: Endothelial
US-07-908-245-1

Query Match 76.5%; Score 13; DB 1; Length 4089;
Best Local Similarity 76.5%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNA 17
|||:|||||
Db 3840 GACTTAGGTGGGAAA 3824

RESULT 10
US-09-123-708-5/c
; Sequence 5, Application US/09123708
; Patent No. 6146887
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Juergen
; APPLICANT: GODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN GENE THERAPEUTIC
; TITLE OF INVENTION: TREATMENT OF VASCULAR DISORDERS
; FILE REFERENCE: 511169-2003
; CURRENT APPLICATION NUMBER: US/09/123,708
; CURRENT FILING DATE: 1998-07-28
; EARLIER APPLICATION NUMBER: 08/553,503

;
; EARLIER FILING DATE: 1996-03-01
; EARLIER APPLICATION NUMBER: P4411402.8
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 5
; LENGTH: 4097
; TYPE: DNA
; ORGANISM: Cytomegalovirus
US-09-123-708-5

Query Match 76.5%; Score 13; DB 3; Length 4097;
Best Local Similarity 76.5%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:| || | |||| |

DB 3868 GACTTAGGTGGGAAA 3852

RESULT 11

US-09-123-624-5/C
; Sequence 5, Application US/09123624
; Patent No. 6149936
; GENERAL INFORMATION:
; APPLICANT: SCHRADER, Jurgen
; APPLICANT: GODECKE, Axel
; TITLE OF INVENTION: DNA EXPRESSION VECTORS FOR USE IN THE GENE THERAPEUTIC
; FILE REFERENCE: 511169-2004
; CURRENT APPLICATION NUMBER: US/09/123,624
; CURRENT FILING DATE: 1998-07-28
; PRIOR APPLICATION NUMBER: 08/553,503
; PRIOR FILING DATE: 1996-03-01
; PRIOR APPLICATION NUMBER: 4411402.8
; PRIOR FILING DATE: 1994-03-31
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 5
; LENGTH: 4097
; TYPE: DNA
; ORGANISM: Bos taurus
US-09-123-624-5

Query Match 76.5%; Score 13; DB 3; Length 4097;
Best Local Similarity 76.5%; Pred. No. 85;
Matches 13; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:| || | |||| |

DB 3868 GACTTAGGTGGGAAA 3852

RESULT 12

US-08-822-028-74
; Sequence 74, Application US/08822028
; Patent No. 5993813
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHOLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; MODIFIED ANTIBODIES FOR CANCER TREATMENT
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DUANE C ULMER
; STREET: P.O. BOX 1967
; CITY: MIDLAND

;
; STATE: MICHIGAN
; COUNTRY: USA
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/822,028
; FILING DATE:

CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/040,687

FILING DATE:
ATTORNEY/AGENT INFORMATION:

NAME: ULMER, DUANE C
REGISTRATION NUMBER: 34,941
REFERENCE/DOCKET NUMBER: C-37,075C
TELECOMMUNICATION INFORMATION:

TELEPHONE: (517) 636-8104
INFORMATION FOR SEQ ID NO: 74:

SEQUENCE CHARACTERISTICS:
LENGTH: 508 base pairs

TYPE: nucleic acid
STRANDEDNESS: double

TOPOLOGY: linear
US-08-822-028-74

Query Match 74.1%; Score 12.6; DB 2; Length 508;
Best Local Similarity 70.6%; Pred. No. 1.1e+02;
Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAYTTYGGNTGGGNA 17
||:|:| || | |||| |

DB 25 GACTTTGGCTGGGGCA 4.

RESULT 13

US-08-479-285-74
; Sequence 74, Application US/08479285
; Patent No. 6207815
; GENERAL INFORMATION:
; APPLICANT: MEZES, PETER S
; APPLICANT: GOURLIE, BRIAN B
; APPLICANT: RIXON, MARK W
; APPLICANT: ANDERSON, WH KERR
; APPLICANT: KAPLAN, DONALD A
; APPLICANT: SCHOLOM, JEFFREY
; TITLE OF INVENTION: A NOVEL FAMILY OF HIGH AFFINITY,
; MODIFIED ANTIBODIES FOR CANCER TREATMENT
; NUMBER OF SEQUENCES: 74
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DUANE C ULMER
; STREET: P.O. BOX 1967
; CITY: MIDLAND
; STATE: MICHIGAN
; COUNTRY: USA
; ZIP: 48641-1967
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/479,285
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 536
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/040687
; FILING DATE: 31-MAR-1993
; ATTORNEY/AGENT INFORMATION:

NAME: ULMER, DUANE C
 REGISTRATION NUMBER: 34,941
 REFERENCE/DOCKET NUMBER: C-37,075C
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (517) 636-8104
 INFORMATION FOR SEQ ID NO: 74:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 508 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-479-285-74

Query Match 74.1%; Score 12.6; DB 4; Length 508;
 Best Local Similarity 70.6%; Pred. No. 1.1e+02;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 25 GACTTTGGCTGGGGCA 41

RESULT 14
 US-08-599-171A-24
 ; Sequence 24, Application US/08599171A
 ; Patent No. 5814473
 ; GENERAL INFORMATION:
 ; APPLICANT: WARREN, Patrick V.
 ; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
 ; NUMBER OF SEQUENCES: 32
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: CARELIA, BYRNE, BAIN, GILFILLAN,
 ; ADDRESSEE: CECCHI, STEWART & OLSTEIN
 ; STREET: 6 BECKER FARM ROAD
 ; CITY: ROSELAND
 ; STATE: NEW JERSEY
 ; COUNTRY: USA
 ; ZIP: 07068
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: 3.5 INCH DISKETTE
 ; COMPUTER: IBM PS/2
 ; OPERATING SYSTEM: MS-DOS
 ; SOFTWARE: WORD PERFECT 5.1
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/599.171A
 ; FILING DATE: Concurrently
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER:
 ; FILING DATE:
 ; CLASSIFICATION:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: HERRON, CHARLES J.
 ; REGISTRATION NUMBER: 28,019
 ; REFERENCE/DOCKET NUMBER: 331400-38
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 201-994-1700
 ; TELEFAX: 201-994-1744
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 912 NUCLEOTIDES
 ; TYPE: NUCLEIC ACID
 ; STRANDEDNESS: SINGLE
 ; TOPOLOGY: LINEAR
 ; MOLECULE TYPE: GENOMIC DNA
 ; US-08-599-171A-24

Query Match 74.1%; Score 12.6; DB 1; Length 912;
 Best Local Similarity 70.6%; Pred. No. 1.2e+02;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 686 GACTTCGGGTGGAGAA 702

RESULT 15
 US-08-646-590B-24
 ; Sequence 24, Application US/08646590B
 ; Patent No. 5962283
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Patrick V.
 ; APPLICANT: Swanson, Ronald V.
 ; TITLE OF INVENTION: TRANSAMINASES AND AMINOTRANSFERASES
 ; NUMBER OF SEQUENCES: 42
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson, P.C.
 ; STREET: 4225 Executive Square, Suite 1400
 ; CITY: La Jolla
 ; STATE: CA
 ; COUNTRY: US
 ; ZIP: 92037
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSEQ for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/646.590B
 ; FILING DATE: 08-May-1996
 ; CLASSIFICATION: 435
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/599,171
 ; FILING DATE: 09-FEB-1996
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: PCT/US97/01094
 ; FILING DATE: 21-January-1997
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Haile, Ph.D., Lisa A.
 ; REGISTRATION NUMBER: 38,347
 ; REFERENCE/DOCKET NUMBER: 09010/017001
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 619/678-5070
 ; TELEFAX: 619/678-5099
 ; INFORMATION FOR SEQ ID NO: 24:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 912 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: single
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: Genomic DNA
 ; FEATURE:
 ; NAME/KEY: Coding Sequence
 ; LOCATION: 1...909
 ; US-08-646-590B-24

Query Match 74.1%; Score 12.6; DB 2; Length 912;
 Best Local Similarity 70.6%; Pred. No. 1.2e+02;
 Matches 12; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Qy 1 GAYTTYGGTGGGNAA 17
 ||:|||||
 Db 686 GACTTCGGGTGGAGAA 702

Search completed: November 5, 2001, 18:05:11
 Job time: 15111 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 5, 2001, 18:01:04 : Search time 5816.79 Seconds
(without alignments)
27.627 Million cell updates/sec

Title: US-08-894-356c-22

Perfect score: 17

Sequence: 1 GAYTYGGNTGGGNAA 17

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 10228115 seqs, 4726426750 residues

Total number of hits satisfying chosen parameters: 20456230

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

- 1: gb_est1:*
- 2: gb_est2:*
- 3: gb_est3:*
- 4: gb_est4:*
- 5: gb_est5:*
- 6: gb_est6:*
- 7: gb_est7:*
- 8: gb_est8:*
- 9: gb_est9:*
- 10: gb_est10:*
- 11: gb_est11:*
- 12: gb_est12:*
- 13: gb_est13:*
- 14: gb_est14:*
- 15: gb_est15:*
- 16: gb_est16:*
- 17: gb_est17:*
- 18: gb_est18:*
- 19: gb_est19:*
- 20: gb_est20:*
- 21: gb_est21:*
- 22: gb_est22:*
- 23: gb_est23:*
- 24: gb_est24:*
- 25: gb_est33:*
- 26: gb_est34:*
- 27: gb_est35:*
- 28: gb_est36:*
- 29: gb_est37:*
- 30: gb_est38:*
- 31: gb_est39:*
- 32: gb_est40:*
- 33: em_estba:*
- 34: em_estfun:*
- 35: em_esthum1:*
- 36: em_esthum2:*
- 37: em_esthum3:*
- 38: em_esthum4:*
- 39: em_esthum5:*
- 40: em_esthum6:*
- 41: em_esthum7:*
- 42: em_esthum8:*
- 43: em_esthum9:*

- 44: em_esthum10:*
- 45: em_esthum11:*
- 46: em_esthum12:*
- 47: em_esthum13:*
- 48: em_esthum14:*
- 49: em_esthum15:*
- 50: em_esthum16:*
- 51: em_esthum17:*
- 52: em_esthum18:*
- 53: em_esthum19:*
- 54: em_esthum20:*
- 55: em_esthum21:*
- 56: em_esthum22:*
- 57: em_esthum23:*
- 58: em_esthum24:*
- 59: em_esthum25:*
- 60: em_esthum26:*
- 61: em_esthum27:*
- 62: em_esthum28:*
- 63: em_estin1:*
- 64: em_estin2:*
- 65: em_estin3:*
- 66: em_estin4:*
- 67: em_estin5:*
- 68: em_estom1:*
- 69: em_estom2:*
- 70: em_estov1:*
- 71: em_estov2:*
- 72: em_estpl1:*
- 73: em_estpl2:*
- 74: em_estpl3:*
- 75: em_estpl4:*
- 76: em_estpl5:*
- 77: em_estpl6:*
- 78: em_estpl7:*
- 79: em_estpl8:*
- 80: em_estpl9:*
- 81: em_estpl10:*
- 82: em_estroi:*
- 83: em_estroi:*
- 84: em_estroi:*
- 85: em_estroi:*
- 86: em_estroi:*
- 87: em_estroi:*
- 88: em_estroi:*
- 89: em_estroi:*
- 90: em_estroi:*
- 91: em_estroi:*
- 92: em_estroi:*
- 93: em_estroi:*
- 94: em_estroi:*
- 95: em_estroi:*
- 96: em_estroi:*
- 97: em_estroi:*
- 98: em_estroi:*
- 99: em_estroi:*
- 100: em_estroi:*
- 101: em_estroi:*
- 102: gb_est25:*
- 103: gb_est26:*
- 104: gb_est27:*
- 105: gb_est28:*
- 106: gb_est29:*
- 107: gb_est30:*
- 108: gb_est31:*
- 109: gb_est32:*
- 110: gb_est41:*
- 111: gb_est42:*
- 112: gb_est43:*
- 113: gb_est44:*
- 114: gb_est45:*
- 115: gb_est46:*
- 116: gb_est47:*

117: gb_est48:*
118: gb_est49:*
119: gb_est50:*
120: gb_est51:*
121: gb_est52:*
122: gb_est53:*
123: gb_est54:*
124: gb_est55:*
125: gb_est56:*
126: gb_est57:*
127: gb_est58:*
128: gb_est59:*
129: gb_est60:*
130: gb_est61:*
131: gb_est62:*
132: gb_est63:*
133: gb_est64:*
134: gb_est65:*
135: gb_est66:*
136: gb_est67:*
137: gb_est68:*
138: gb_est69:*
139: gb_est70:*
140: gb_est71:*
141: gb_est72:*
142: gb_est73:*
143: gb_est74:*
144: gb_est75:*
145: gb_est76:*
146: gb_est77:*
147: gb_est78:*
148: gb_est79:*
149: gb_est80:*
150: gb_est81:*
151: gb_est82:*
152: gb_est83:*
153: gb_est84:*
154: gb_est85:*
155: gb_est86:*
156: gb_est87:*
157: gb_est88:*
158: gb_est89:*
159: gb_est90:*
160: gb_est91:*
161: gb_est92:*
162: gb_est93:*
163: gb_est94:*
164: gb_est95:*
165: gb_est96:*
166: gb_est97:*
167: gb_est98:*
168: gb_est99:*
169: gb_est100:*
170: gb_est101:*
171: gb_est102:*
172: gb_est103:*
173: gb_est104:*
174: gb_est105:*
175: gb_est106:*
176: gb_est107:*
177: gb_est108:*
178: gb_est109:*
179: gb_est110:*
180: gb_est111:*
181: gb_est112:*
182: gb_est113:*
183: gb_est114:*
184: gb_est115:*
185: gb_est116:*
186: gb_est117:*
187: gb_est118:*
188: gb_est119:*
189: gb_est120:*

190: gb_est121:*
191: gb_est122:*
192: gb_est123:*
193: gb_est124:*
194: gb_est125:*
195: gb_est126:*
196: gb_est127:*
197: gb_est128:*
198: gb_est129:*
199: gb_est130:*
200: gb_est131:*
201: gb_est132:*
202: gb_est133:*
203: gb_est134:*
204: gb_est135:*
205: gb_est136:*
206: gb_est137:*
207: gb_est138:*
208: gb_est139:*
209: gb_est140:*
210: gb_est141:*
211: gb_est142:*
212: gb_est143:*
213: gb_est144:*
214: gb_est145:*
215: gb_est146:*
216: gb_est147:*
217: gb_est148:*
218: gb_est149:*
219: gb_est150:*
220: gb_est151:*
221: gb_est152:*
222: gb_est153:*
223: gb_est154:*
224: gb_est155:*
225: gb_est156:*
226: gb_est157:*
227: gb_est158:*
228: gb_est159:*
229: gb_est160:*
230: gb_est161:*
231: gb_est162:*
232: gb_est163:*
233: gb_est164:*
234: gb_est165:*
235: gb_est166:*
236: gb_est167:*
237: gb_est168:*
238: gb_est169:*
239: gb_est170:*
240: gb_est171:*
241: gb_est172:*
242: gb_est173:*
243: gb_est174:*
244: gb_est175:*
245: gb_est176:*
246: gb_est177:*
247: gb_est178:*
248: gb_est179:*
249: gb_est180:*
250: gb_est181:*
251: gb_est182:*
252: gb_est183:*
253: gb_est184:*
254: gb_est185:*
255: gb_est186:*
256: gb_est187:*
257: gb_est188:*
258: gb_est189:*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

AW234169 244 bp mRNA EST 17-JUL-2000
 LOCUS sf22a08.y1 Gm-cl028 Glycine max cDNA clone GENOME SYSTEMS CLONE ID:
 DEFINITION Gm-cl028-687 5' similar to TR:Q43583 Q43583 HSR201 PROTEIN. ;, mRNA
 sequence.
 ACCESSION AW234169
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM soybean.
 Glycine max
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.
 REFERENCE 1 (bases 1 to 244)
 AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Corvett,V., Khanna
 A., Bolla,B., Marr,M., Hillier,L., Kucaba,T., Martin,J., Beck,C.,
 Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers
 Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk
 R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann
 R., Waterston,R. and Wilson,R.
 TITLE Public Soybean EST Project
 JOURNAL
 COMMENT Unpublished (1999)
 Contact: Shoemaker R/Public Soybean EST Project
 Public Soybean EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 This clone is available through: Genome Systems, Inc. 4633 World
 Parkway Circle St. Louis, Missouri 63134 For further information
 call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314)
 427-3324 or contact: clones@genomesystems.com or
 info@genomesystems.com web site: www.genomesystems.com
 Insert Length: 486 Std Error: 0.00.

FEATURES

Location/Qualifiers
 1..244
 /organism="Glycine max"
 /db_xref="taxon:3847"
 /clone="GENOME SYSTEMS CLONE ID: Gm-cl028-687"
 /clone_lib="Gm-cl028"
 /tissue_type="roots of 'Supernod' plants"
 /lab_host="DH10B"
 /note="vector: pBluescript II XR; Site_1: EcoRI; Site_2:
 XhoI; The mRNA was isolated from roots of Glycine max
 'Supernod' plants generously donated by Dr. Gary Stacey.
 The seedlings were inoculated with Bradyrhizobium
 japonicus, strain USDA110 prior to harvest. Stratagene's
 cDNA synthesis kit (catalog number 200401) was used to
 synthesize the cDNA. First-strand synthesis was performed
 with 5-methyl dCTP, hence the ligated cDNA was
 hemimethylated. A modification of Stratagene's
 first-strand synthesis primer was used. An 'anchor'
 nucleotide (V-A-C, or G) was added to the 3' end of the
 primer (GAGAGAGAGAGAGAGAGACTAGTCGAG(T)18V) to anchor
 the primer at the 5' end of the poly(A) tract. After
 second-strand synthesis, the cDNA ends were filled in with
 cloned Pfu DNA polymerase, ligated to EcoRI adapters and
 subsequently phosphorylated. The XhoI site within the
 first-strand synthesis primer was then restricted by
 digestion with XhoI; all XhoI sites in the cDNA would be
 protected by their hemimethylated status. The cDNA
 constructs were size-fractionated with a 500bp cutoff,
 using GibcoBRL Life Technologies' cDNA size fractionation
 column. The column eluent was then ligated into
 Stratagene's pBluescript II XR Predigested vector
 (pBluescript II SK(+)) that has been digested with EcoRI
 and XhoI, and phosphorylated by Stratagene. Both the
 white and blue colonies appear to contain recombinant
 plasmids with cDNA inserts, based on size (n=25). This
 library was constructed by Dr. Paul Keim and Dr. Virginia
 Corvett."

65 a 37 c 65 g 77 t

ORIGIN

Query Match 83.5%; Score 14.2; DB 113; Length 244;
 Best Local Similarity 76.5%; Pred. No. 1.2e+03;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNA 17
 ||:|:|:|:|:|:|:|:|:|
 Db 10 GACTTTGGTTGGGGAA 26

RESULT 3

LOCUS AI488429/c
 DEFINITION EST245768 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
 CLEDD21D13, mRNA sequence.
 ACCESSION AI488429
 VERSION
 KEYWORDS EST.
 SOURCE tomato.
 ORGANISM Lycopersicon esculentum
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Asteridae; euasterids I; Solanales; Solanaceae; Solanum;

REFERENCE 1 (bases 1 to 247)

Aucala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
 Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman
 C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley
 S.D. and Giovannoni,J.
 Generation of ESTs from tomato carpel tissue
 Unpublished (1999)
 Contact: David Frisch
 Clemson University Genomics Institute
 Clemson University
 100 Jordan Hall, Clemson, SC 29634, USA
 Tel: 864 656 4366
 Fax: 864 656 4293
 Email: dfrisch@clemson.edu.

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers
 1..247
 /organism="Lycopersicon esculentum"
 /cultivar="TA496"
 /db_xref="taxon:4081"
 /clone="CLEDD21D13"
 /clone_lib="tomato ovary, TAMU"
 /tissue_type="carpel"
 /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
 /lab_host="XLI-Blue MRF"
 /note="vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
 XhoI; cLED - Tomato Carpel EST Library_ OligodT-primed and
 directionally cloned cDNA in vector Lambda ZAP II with 5'
 and 3' ends located at the EcoRI and XhoI sites,
 respectively."

BASE COUNT 72 a 67 c 27 g 81 t

ORIGIN

Query Match 83.5%; Score 14.2; DB 21; Length 247;
 Best Local Similarity 76.5%; Pred. No. 1.2e+03;
 Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNA 17
 ||:|:|:|:|:|:|:|:|:|
 Db 222 GATTGTGGATGGGAA 206

RESULT 4

LOCUS BG550470
 DEFINITION BG550470 247 bp mRNA EST 09-APR-2001
 947076G02.x2 947 - 2 week shoot from Barkan lab zea mays cDNA, mRNA
 sequence.

```

ACCESSION      BG550470
VERSION         BG550470.1  GI:13562250
KEYWORDS        EST.
SOURCE          Zea mays.
ORGANISM        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACC
                clade; Panicoideae; Andropogoneae; Zea.
REFERENCE       1 (bases 1 to 247)
AUTHORS         Walbot V.
TITLE           Maize ESTs from various cDNA libraries sequenced at Stanford
                University
JOURNAL         Unpublished (1999)
COMMENT         Contact: Walbot V
                Department of Biological Sciences
                Stanford University
                855 California Ave, Palo Alto, CA 94304, USA
                Tel: 650 723 2227
                Fax: 650 725 8221
                Email: walbot@stanford.edu
                Plate: 947076 row: G column: 02.
FEATURES       Location/Qualifiers
                1..247
                /organism="Zea mays"
                /cultivar="B73"
                /db_xref="taxon:4577"
                /clone_lib="947 - 2 week shoot from Barkan lab"
                /tissue_type="leaf and stem, including leaf base"
                /dev_stage="2 week old seedling (3 leaves)"
                /lab_host="X11-Blue"
                /note="Organ: shoot; Vector: Lambda ZAP (pBlueScript SK-);
                Site:1: EcoRI; Site:2: XhoI; Directionally cloned using
                Stratagene's Unizap XR cDNA cloning kit with the 5' end
                at the EcoRI site. The library represents 8 x 105
                independent recombinant phage. The plants were greenhouse
                grown."
BASE COUNT      57 a 63 c 85 g 42 t
ORIGIN
Query Match      83.5%; Score 14.2; DB 155; Length 247;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTGGNTGGGNA 17
   ||:|||||
Db 224 GATTTGGTGGGAAA 240

RESULT 5
LOCUS      AZ015624
DEFINITION RPI-23-258D7.TJ RPI-23 Mus musculus genomic clone RPI-23-258D7,
DNA sequence.
ACCESSION  AZ015624
VERSION     AZ015624.1  GI:7091008
KEYWORDS    GSS.
SOURCE      house musculus.
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1 (bases 1 to 249)
AUTHORS     Zhao,S., Nierman,W., Feldblyum,T., Malek,J., Shatsman,S., Aklnret
            ,B., Levins,M., McGann,S., Tsegaye,G., Geer,K., Krol,M., de Jong,P.
            and Fraser,C.M.
TITLE       Mouse BAC End Sequences from Library RPI-23
JOURNAL     Unpublished (1999)
COMMENT     Contact: Shaying Zhao
            Department of Eukaryotic Genomics
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200

```

```

Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/orderingframe.htm)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bac_ends/mouse/bac_end_intro.html
Plate: 258 row: D column: 7
Seq primer: SP6
Class: BAC ends.
FEATURES       Location/Qualifiers
                1..249
                /organism="Mus musculus"
                /strain="C57BL/6J"
                /db_xref="taxon:10090"
                /clone_lib="RPI-23-258D7"
                /clone_lib="RPI-23"
                /sex="Female"
                /lab_host="DH10B"
                /note="Organ: Kidney/Brain; Vector: pBACe3.6; Site:1:
                EcoRI; Site:2: EcoRI; Female C57BL/6J mouse kidney and/or
                brain genomic DNA was isolated and partially digested
                with a combination of EcoRI and EcoRI Methylase. Size
                selected DNA was cloned into the pBACe3.6 vector at the
                EcoRI sites. The ligation products were transformed into
                DH10B electrocompetent cells (BRL Life Technologies)."
BASE COUNT      87 a 54 c 39 g 69 t
ORIGIN
Query Match      83.3%; Score 14.2; DB 237; Length 249;
Best Local Similarity 76.3%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 GAYTTGGNTGGGNA 17
   ||:|||||
Db 60 GATTTGGTGGGCAA 76

RESULT 6
LOCUS      AA650754
DEFINITION 30868 Lambda-PRL2 Arabidopsis thaliana cDNA clone 281H6T7, mRNA
sequence.
ACCESSION  AA650754
VERSION     AA650754.1  GI:2580846
KEYWORDS    EST.
SOURCE      thale cress.
ORGANISM    Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE   1 (bases 1 to 258)
AUTHORS     Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh
            ,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomashow,M., Retzel
            ,E. and Somerville,C.
TITLE       Genes galore: a summary of methods for accessing results from
            large-scale partial sequencing of anonymous Arabidopsis cDNA clones
            Plant Physiol. 106. 1241-1255 (1994)
JOURNAL     95148729
MEDLINE
COMMENT     Contact: Thomas Newman
            MSU-DOE Plant Research Laboratory
            Michigan State University
            MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
            Lansing, MI
            Tel: 517-353-0854
            Fax: 517-353-9168
            Email: 22313tcm@lbn.cl.msu.edu
            Seq primer: T7 dye primer.
            Location/Qualifiers
                1..258
                /organism="Arabidopsis thaliana"

```

/strain="var columbia"
/db_xref="taxon:3702"
/clone="281H67"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a CDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda Zip-Lox. The CDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

BASE COUNT 74 a 42 c 62 g 70 t 10 others
ORIGIN

Query Match 83.5%; Score 14.2; DB 10; Length 258;
Best Local Similarity 76.5%; Pred. NO. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GAYTTYGGNTGGGNAA 17
||:||||| ||||| ||
Db 53 GATTCGGTGGGGTAA 69

RESULT 7
BG046171
LOCUS
DEFINITION BG046171 259 bp mRNA EST 25-JAN-2001
ID: Gm-c1060-215 5' similar to TR:Q43583 Q43583 HSR201 PROTEIN. ; mRNA sequence.

ACCESSION BG046171
VERSION BG046171.1 GI:12494655
KEYWORDS EST.
SOURCE soybean.

ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE 1 (bases 1 to 259)
AUTHORS Shoemaker,R., Kelim,P., Vodkin,L., Erpelting,J., Coryell,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX:(888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 157.

FEATURES
source
1. 259
Location/Qualifiers
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1060-215"
/clone_lib="Gm-c1060"
/tissue_type="Root, 2 week seedlings"
/lab_host="DH10B"
/note="Vector: pBluescript II SK+; Site_1: EcoRI; Site_2: XhoI; The CDNA library was constructed from mRNA isolated

from root tissue of 2 week old seedlings for PI468916.
Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with a XhoI restriction site. EcoRI adapters were ligated to the blunt-ended CDNA fragments followed by XhoI digestion. The CDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated CDNA fragments were transformed into DH10B host cells (Gibco BRL). This library was constructed in the laboratory of Dr. Randy Shoemaker at Iowa state university."

BASE COUNT 50 a 41 c 91 g 77 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 173; Length 259;
Best Local Similarity 76.5%; Pred. NO. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GAYTTYGGNTGGGNAA 17
||:||||| ||||| ||
Db 89 GACTTTGGTGGGGGAA 105

RESULT 8
AA651035/c
LOCUS
DEFINITION AA651035 263 bp mRNA EST 05-JAN-1998
ID: 31149 Lambda-PRL2 Arabidopsis thaliana CDNA clone 168A5XP 3', mRNA sequence.

ACCESSION AA651035
VERSION AA651035.1 GI:2581127
KEYWORDS EST.
SOURCE thale cress.

ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsids.

REFERENCE 1 (bases 1 to 263)
AUTHORS Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H., McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M., Retzel,E. and Somerville,C.
Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cDNA clones Plant Physiol. 106, 1241-1255 (1994)
95148729

TITLE
JOURNAL
MEDLINE
COMMENT Contact: Thomas Newman
MSU-DOE Plant Research Laboratory
Michigan State University
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI
Tel: 517-353-0854
Fax: 517-353-9168
Email: 22313t@icbm.cl.msu.edu
The sequence entry for this EST has been reverse complemented and is being submitted in the sense orientation.
Seq primer: M13_Universal.

FEATURES
source
1. 263
Location/Qualifiers
/organism="Arabidopsis thaliana"
/strain="var columbia"
/db_xref="taxon:3702"
/clone="168A5XP"
/clone_lib="Lambda-PRL2"
/note="Vector: lambda Zip-Lox; Site_1: Sal; Site_2: Not; Lambda PRL2 is a CDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques. The vector is BRL's lambda Zip-Lox. The CDNA inserts were directionally cloned with Sal-Not arms using oligo dt primed cDNA."

BASE COUNT 56 a 68 c 50 g 79 t 10 others

ORIGIN	BB264508/c									
	LOCUS									
Query Match	83.5%; Score 14.2; DB 10; Length 263;									
	Best Local Similarity 76.5%; Pred. No. 1.2e+03;									
Matches	13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;									
QY	1	GAYTTYGGNTGGGNA 17								
		:								
Db	90	GATTCGGATGGGCCA 74								
RESULT	9									
	AI608251/c									
LOCUS	278 bp mRNA EST									
	21-APR-1999									
DEFINITION	v93f09 x1 Knowles Solter mouse embryonic stem cell Mus musculus									
	CDNA clone IMAGE:894569 3', mRNA sequence.									
ACCESSION	AI608251									
	AI608251									
VERSION	AI608251.1									
	GI:4617418									
KEYWORDS	EST.									
	EST.									
SOURCE	house mouse.									
	Mus musculus									
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.									
REFERENCE	1 (bases 1 to 278)									
	Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.									
AUTHORS	The WashU-NCI Mouse EST Project 1999									
	Unpublished (1999)									
TITLE	Contact: Marra M/WashU-NCI Mouse EST Project 1999									
	Washington University School of Medicine									
JOURNAL	4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA									
	Tel: 314 286 1800									
COMMENT	Fax: 314 286 1810									
	Email: mouseest@wustl.edu									
FEATURES	This clone is available royalty-free through LLNL: contact the IMAGE Consortium (info@image.llnl.gov) for further information.									
	MGI:522529									
source	This clone was previously sequenced on the 5' end only, this new data is from the 3' end									
	Seq primer: Primer name ambiguous									
Location/Qualifiers	High quality sequence stop: 274.									
	1..278									
source	/organism="Mus musculus"									
	/strain="B6D2 F1/J"									
source	/db_xref="taxon:10090"									
	/clone="IMAGE:894569"									
source	/clone_lib="Knowles Solter mouse embryonic stem cell"									
	/dev_stage="embryo"									
source	/lab_host="DH10B"									
	/note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; Cloned unidirectionally from mRNA prepared from 800 blastocysts. Primer: SalI(dT): 5'-CGGTCGACCGTCGACCGTTTTTTTTTTT-3'.									
source	CDNAs were cloned into the NotI/SalI sites of a pSPORT vector (Life Technologies).									
	vector (Life Technologies).									
BASE COUNT	57 a 75 c 48 g 98 t									
ORIGIN	83.5%; Score 14.2; DB 22; Length 278;									
	Best Local Similarity 76.5%; Pred. No. 1.2e+03;									
Matches	13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;									
QY	1	GAYTTYGGNTGGGNA 17								
		:								
Db	99	GATTTGGTGGGCCAA 83								
RESULT	10									

BB264508 294 bp mRNA EST 07-JUL-2000
 BB264508 RIKEN full-length enriched, 10 days neonate cortex Mus
 musculus CDNA clone A830015G22 3', mRNA sequence.
 BB264508
 BB264508.1 GI:8960965
 EST.
 house mouse.
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 294)
 Konno, H., Aizawa, K., Akahira, S., Akiyama, J., Arakawa, T., Carninci
 P., Endo, T., Fukuda, S., Fukunishi, Y., Hara, A., Hayatsu, N.,
 Hirozane, T., Hori, F., Ishii, Y., Ishikawa, J., Ishikawa, T., Itoh, M.,
 Izawa, M., Kadota, K., Kagawa, I., Kai, C., Kawai, J., Kikuchi, N.,
 Kiyosawa, H., Kojima, Y., Kondo, S., Koya, S., Kurlhara, C., Kusakabe, M.,
 Matsuyama, T., Miki, R., Mizuno, Y., Nakamura, M., Oda, H., Okazaki, Y.,
 Ono, T., Owa, C., Saito, H., Sakai, C., Sato, K., Shibata, K., Shibata, Y.,
 Y., Shigemoto, Y., Shinagawa, A., Shiraki, T., Sogabe, Y., Sugahara, Y.,
 Suzuki, H., Suzuki, H., Tagawa, A., Takahashi, F., Tominaga, N., Toya
 T., Tsunoda, Y., Watahiki, A., Watanabe, S., Yamamura, T., Yamanaka, I.,
 Yano, K., Yasunishi, A., Yokota, T., Yoshida, K., Yoshiki, A., Yoshino
 M., Muramatsu, M. and Hayashizaki, Y.
 RIKEN Mouse ESTs (Konno, H., et al.)
 Unpublished (2000)
 Contact: Yoshihide Hayashizaki
 Genome Exploration Research Group, Life Science Tsukuba Center,
 Genome Science Laboratory
 The Institute of Physical and Chemical Research (RIKEN), Genomic
 Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: +81-298-36-9013
 Fax: +81-298-36-9098
 Email: genome-res@rtc.riken.go.jp/
 URL: http://genome.rtc.riken.go.jp/
 Carninci, P., Nishiyama, Y., Westover, A., Itoh, M., Nagaoaka, S., Sasaki
 N., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Thermostabilization and thermoactivation of thermolabile enzymes by
 trehalose and its application for the synthesis of full length
 cDNA. Proc. Natl. Acad. Sci. U.S.A. 95 (2), 520-524 (1998)
 Itoh, M., Kitsuina, T., Akiyama, J., Shibata, K., Izawa, M., Kawai, J.,
 Tomaru, Y., Carninci, P., Shibata, Y., Ozawa, Y., Muramatsu, M., Okazaki
 Y. and Hayashizaki, Y.
 Automated filtration-based high-throughput plasmid preparation
 system. Genome Res. 9 (5), 463-470 (1999)
 Carninci, P. and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning. Methods Enzymol. 303,
 19-44 (1999)
 Please visit our web site (http://genome.rtc.riken.go.jp) for
 further details.
 Location/Qualifiers
 1..294
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="A830015G22"
 /clone_lib="RIKEN full-length enriched, 10 days neonate
 cortex"
 /tissue_type="cortex"
 /dev_stage="10 days neonate"
 /lab_host="DH10B"
 /note="Site_1: SalI; Site_2: BamHI; cDNA library was
 prepared and sequenced in Mouse Genome Encyclopedia
 Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in
 RIKEN. Division of Experimental Animal Research in Riken
 contributed to prepare mouse tissues. 1st strand cDNA was
 primed with a primer [5'
 GAGAGAGAGAGATCCAGAGACTCTTTTTTTTTTTTNN 3']. cDNA was
 prepared by using trehalose thermo-activated reverse
 transcriptase and subsequently enriched for full-length by
 cap-trapper. cDNA went through one round of normalization
 to Rot = 20.0 and subtraction to Rot = 459.0. Second

strand cDNA was prepared with the primer adapter of sequence [5' GAGAGAGATTCGAGTTAATAATTAATCCCCCCCC 3']. cDNA was cleaved with XhoI and BamHI. Vector: a modified pBluescript KS(+) after bulk excision from Lambda FLC I."

BASE COUNT 64 a 77 c 66 g 87 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 129; Length 294;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNAA 17
||:|:|:|:|:|:|:|:|:|
Db 243 GACTTTGGTGGGGAA 227

RESULT 11

AM471629
LOCUS AM471629 307 bp mRNA EST 24-FEB-2000
DEFINITION sll3h01.v1 Gm-cl029 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl029-1130 5' similar to TR:004201 004201
HYPERSENSITIVITY-RELATED GENE 201 ISOLOG. ;, mRNA sequence.

ACCESSION AM471629
VERSION AM471629.1 GI:7041735
KEYWORDS EST.
SOURCE soybean.

ORGANISM

Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE

1 (bases 1 to 307)
Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Corvett,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
High quality sequence stop: 301.

FEATURES

source

1..307
Location/Qualifiers
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl029-1130"
/clone_lib="Gm-cl029"
/tissue_type="very young cotyledons of greenhouse grown plants"
/lab_host="DH10B"
/note="Vector: pSPORT1; Site_1: Not I; Site_2: Sal I; This cDNA library was constructed from mRNA isolated from very young cotyledons (20-50mgs fresh weight) of greenhouse grown plants. The library was prepared using the Life Technologies pSuperScript cDNA library construction kit. Complementary DNA was synthesized from mRNA using a poly(dT) sequence with a Not I restriction site. Sal I linker adapters were ligated to the blunt-ended cDNA fragments followed by Not I digestion. The cDNA fragments were directionally cloned into the Not I-Sal I restriction

site of the pSPORT1 vector. The ligated cDNA fragments were transformed into E.coli Electromax DH10B host cells. This library was constructed by Dr. Paul Keim and Dr. Virginia Corvett."

BASE COUNT 96 a 33 c 96 g 82 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 116; Length 307;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAYTTYGGNTGGGNAA 17
||:|:|:|:|:|:|:|:|:|
Db 127 GATTTGGATGGGGAA 143

RESULT 12

AI856623
LOCUS AI856623 324 bp mRNA EST 16-JUL-1999
DEFINITION sb39g12.v1 Gm-cl014 Glycine max cDNA clone GENOME SYSTEMS CLONE ID: Gm-cl014-95 5' similar to TR:004201 004201
HYPERSENSITIVITY-RELATED GENE 201 ISOLOG. ;, mRNA sequence.

ACCESSION AI856623
VERSION AI856623.1 GI:5510239
KEYWORDS EST.
SOURCE soybean.

ORGANISM

Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Glycine.

REFERENCE

AUTHORS

1 (bases 1 to 324)
Shoemaker,R., Keim,P., Vodkin,L., Erpeliding,J., Corvett,V., Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J., Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M., Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N., Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,Y., Cardenas,M., McCann,R., Waterston,R. and Wilson,R.
Public Soybean EST Project
Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@watson.wustl.edu
This clone is available through: Genome Systems, Inc. 4633 World Parkway Circle St. Louis, Missouri 63134 For further information call: (800) 430-0030 or (314) 427-3222 FAX: (888) 919-3324 or (314) 427-3324 or contact: clones@genomesystems.com or info@genomesystems.com web site: www.genomesystems.com
Seq primer: -40RP from Gbco
High quality sequence stop: 311.

FEATURES

source

1..324
Location/Qualifiers
/organism="Glycine max"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-cl014-95"
/clone_lib="Gm-cl014"
/tissue_type="Leaves, 2-3 week old seedlings, greenhouse grown"
/lab_host="DH10B"
/note="Vector: pT73Pac (pT73, Pharmacia); Site_1: EcoRI; Site_2: HindIII; This cDNA library was constructed from mRNA isolated from leaves of 2-3 week old greenhouse grown plants. Complementary DNA was synthesized from mRNA using a 3' anchored poly (dT) primer. EcoRI adapters were ligated to the blunt-ended cDNA fragments followed by digestion with EcoRI and HindIII. The cDNA fragments were directionally cloned into the EcoRI-HindIII restriction site of the pT73-Pac vector. The ligated cDNA fragments

RESULT 15
BE460581/C

LOCUS BE460581 347 bp mRNA EST 27-JUL-2000
DEFINITION EST412000 tomato breaker fruit, TIGR Lycopersicon esculentum cDNA
clone cLEG31N6, mRNA sequence.
ACCESSION BE460581
VERSION BE460581.1 GI:9504883
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 347)
AUTHORS Alcala,J., Vrebalov,J., White,R., van der Hoeven,R.S., Holt,I.E.,
Liang,F., Hansen,T.S., Craven,M.B., Bowman,C.L., Ronning,C.M.,
Nierman,W., Fraser,C.M., Martin,G.B., Giovannoni,J.J. and Tanksley
S.D.
TITLE Generation of ESTs from tomato fruit tissue, breaker stage
JOURNAL Unpublished (2000)
COMMENT Contact: David Frisch
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 4366
Fax: 864 656 4293
Email: dfrisch@clemson.edu
5 prime sequence.
FEATURES
Source
1..347
Location/Qualifiers
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="cLEG31N6"
/clone_lib="tomato breaker fruit, TIGR"
/tissue_type="Pericarp"
/dev_stage="breaker"
/lab_host="SOLR"
/note="Vector: pBluescriptSKmCvadapt; Site.1: EcoRI;
Site.2: XhoI; Fruit were harvested at the breaker stage
(first sign of lycopene accumulation on the blossom end of
the fruit). Fruit were cut in half and the seeds and
locules were discarded prior to freezing the pericarp."
BASE COUNT 120 a 71 c 40 g 116 t
ORIGIN

Query Match 83.5%; Score 14.2; DB 167; Length 347;
Best Local Similarity 76.5%; Pred. No. 1.2e+03;
Matches 13; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Oy 1 GATTTGGTGGGNAA 17
||:|||||
Db 91 GATTTGGTGGGAAA 75
||:|||||

Search completed: November 5, 2001, 18:01:08
Job time: 14908 sec